DOCUMENT RESUME

ED 258 602 JC 840 010

TITLE	Identification and Assessment of Learning Disabled Students at the California Community Colleges: Final Report.
INSTITUTION	Consortium for the Study of Learning Disabilities in the California Community Colleges.
SPONS AGENCY	California Community Colleges, Sacramento. Office of the Chancellor.
PUB DATE	18 Oct 83
NOTE	186p.
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF01/PC08 Plus Postage.
DESCRIPTORS	*Ancillary School Services; Community Colleges; Diagnostic Tests; Eligibility; *Handicap Identification; *Learning Disabilities; State Surveys; Two Year Colleges; *Two Year College Students
IDENTIFIERS	*California

ABSTRACT

This report outlines the initial results of a study concerning the procedures and criteria used to identify learning disabled (LD) students and eligibility criteria of community college LD programs in California as mandated by the state legislature. First, an introductory section discusses the purposes of the report, background on LD research and guidelines, the initiation and purpose of the research project, project participants, and the problems addressed by the project. An overview of California community college LD programs is followed by a discussion of project goals and expected outcomes, indicating that the four project goals were: (1) to specify educationally relevant attributes and technically adequate instruments that differentiate learning disabled adults from other low-achieving adults; (2) to develop appropriate discrepancy/selection models to certify the presence of a learning disability in community college adults; (3) to determine the magnitude of discrepancy and/or cutoff levels necessary to verify a learning disability among community college students; and (4) to empirically validate the verifying properties (i.e., attributes, models, and cutoff levels) of a learning disability. The next sections discuss each of these goals in terms of pertinent activities, problems, resolutions, and products. The final sections provide a summary and recommendations for action. (LAL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED258602

FINAL REPORT

Identification and Assessment of Learning Disabled Students at the California Community Colleges

prepared for

OFFICE OF THE CHANCELLOR
CALIFORNIA COMMUNITY COLLEGES

submitted by

THE CONSORTIUM FOR THE STUDY OF LEARNING DISABILITIES
IN THE CALIFORNIA COMMUNITY COLLEGES

October 18, 1983

JC 840 010

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
G. Hayward
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☒ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

This final report has been prepared by the Steering Committee of the Consortium for the Study of Learning Disabilities in the California Community Colleges.

Ron Baker, Mira Costa Community College
Laurel Best, Cypress Community College
Donald D. Deshler, University of Kansas
Robert Howard, Chancellor's Office
Daryl Mellard, University of Kansas
Bruce Ostertag, California State University at Sacramento

The Consortium wishes to acknowledge the assistance and expertise of the following research assistants and staff of the University of Kansas Institute for Research in Learning Disabilities

Research Assistants	Staff
Sidney Cooley	Muriel Hays
Thurma DeLoach	Peggy Showalter
Kathy Hori	Eleanor Womack
Javaid Kaiser	

TABLE OF CONTENTS

TABLE OF CONTENTS
READER'S GUIDE 1
INTRODUCTION 2

GOAL 1 20
GOAL 2 65
GOAL 3 87
GOAL 4 98

SUMMARY 156
RECOMMENDATIONS

READER'S GUIDE

The complexity of issues and number of pages in this report are such that we offer this guide to assist the reader in the location of specific information. This report has been organized into eight sections. These sections are: (a) an introduction, (b) information related to Goal 1, (c) information related to Goal 2, (d) information related to Goal 3, (e) information related to Goal 4, (f) a summary, and (g) a set of recommendations for proposed action.

The reader may have occasion to need a comparatively quick review of the past year's goals, objectives, procedures and outcomes. This information and problems confronted in the identification of learning disabled students is presented in an abbreviated manner in the Summary section. The additional, specific information can be found in the body of the report.

INTRODUCTION

Purpose of the report. This report outlines the initial results of the study concerning the identification and eligibility criteria of community college learning disabilities (LD) programs as mandated by the legislature in supplemental budget language. In a research project with as broad a scope and with such far-reaching consequences as the development of procedures for identifying learning disabled adults in community college settings, it is appropriate, and perhaps mandatory, to document clearly the underlying process and the decisions from which the recommended identification procedures are evolving. This report provides an inclusive accounting of the activities which resulted in an initial set of procedures for identifying learning disabled students in California's community colleges. At times, the attention afforded specific details and explanations may appear tedious and of little consequence. However, given the exploratory nature of the research and the numerous "fork-in-the-road" decisions that had to be made, we believe that acceptance of our recommendations will depend on a thorough understanding of the procedures followed and the values underlying key decisions.

The following text describes the development of the project from its inception through the end of the first year. The project was organized around four goals, each of which will be discussed in terms of the pertinent activities, problems, resolutions, and products. A comprehensive accounting of a project of this magnitude entails the danger of including so much technically sophisticated data that the nontechnical reader is confused. Therefore, to help insure that the information is readily understandable, the main body of the report contains a nontechnical description of the project, while the specific technical information, task force reports, research instruments, etc., have been organized into a set of appendices. From a researcher's perspective, the process reported here was challenging, frustrating, and exhilarating. From the standpoint of the researcher and practitioner, the products will undoubtedly be recognized as only the beginning of the answer to the problem of reliable identification of LD adults in community college settings.

Background. Recent studies undertaken by the California Department of Finance and the Office of the Chancellor of California's community colleges have revealed considerable variability in the procedures and criteria used to identify learning disabled students within the California community college system. Documentation of these inconsistencies has resulted in growing concern that statewide inequities may exist in: (a) the categorical aid (excess costs) provided by the state to community college LD programs; and (b) the availability of service to students when the eligibility criteria for one LD program can be more or less exclusive than those of another LD program. To rectify this situation, the California legislature included supplemental budget language requiring the development of minimum eligibility standards for assessing and identifying LD adults in the community colleges.

It comes as no surprise that California's community college LD programs lack consistent eligibility criteria. Current federal policy (PL 94-142) bases the identification of a learning disability on the existence of a significant discrepancy between a student's assessed intellectual ability and academic achievement in the absence of any other primary handicapping conditions or disadvantage. According to the Federal Register (December 29, 1977) "when a severe discrepancy exists which cannot be explained by the presence of other known factors that lead to such a discrepancy, the cause is believed to be a specific learning disability" (p. 65085). However, attempts to use PL 94-142's guidelines for identifying LD students have proven problematic. Thus, the concept of ability/achievement discrepancy, though crucial to the definition of learning disabilities, is not easily operationalized, particularly at the community college level. Furthermore, the federal regulations do not adequately address the difficulty of differentiating LD from other low-achieving (LA) students.

In particular, uncertainty surrounds the following aspects of the identification dilemma: how much of an aptitude-achievement discrepancy is necessary, which tests are valid and reliable for adult-aged populations, what abilities and skills should be measured, and what exclusions should apply. Additionally, discrepancy alone does not adequately tap the true nature of a learning disability, i.e., a handicap in the ability to learn--not just a measured difference between what an individual has learned and what he/she is expected to have learned. All of these questions and uncertainties make current inconsistencies in California's community college LD identification procedures inevitable. Indeed, other research has shown similar inconsistencies in LD programs for all ages nationwide.

Due to a lack of clear guidelines, California's community college districts have developed their own criteria for assessing and identifying LD adults. A survey initiated by the Chancellor's Office (Ostertag & Baker, 1983) of LD program characteristics, identification, and assessment practices in California's community colleges demonstrated a general consistency among California's 106 community colleges in the methods used to identify and assess LD (i.e., testing instrument used, referral sources, use of intake interviews).

However, specific criteria for determining eligibility for LD services were not adequately defined and many of the assessment instruments typically used lack sufficient reliability to justify their continued use. Ostertag and Baker recommended that a more refined approach to identification and assessment be implemented and that a system for measuring a student's ability/achievement discrepancy become part of the identification process. Additionally, the authors pointed out that population-appropriate, valid and reliable assessment tools must replace currently used tests of questionable technical adequacy.

Initiation of the project. Laurel Best, Learning Disability Division Chairperson for the California Association of Post-Secondary Education of the Disabled and a member of the Advisory Committee to the Department of Finance Program Review Analysts, in an effort to accommodate Ostertag and Baker's (1983) recommendations, contacted the University of Kansas Institute for Research in Learning Disabilities (KU-IRLD). The KU-IRLD, which was already involved in research focusing on secondary and post-secondary LD populations, was asked to make suggestions for how to resolve the problem of identifying LD students in community college settings. Coincindentally, the KU-IRLD core staff had recently formulated a proposal for distinguishing LD children and youth from other low-achieving (LA) individuals, which addressed many of the concerns articulated by Ostertag and Baker. Consequently, this proposal was presented to the Chancellor's Office of the California community colleges as a prototype for developing procedures for identifying and assessing LD community college students. The Chancellor's Office offered the KU-IRLD a contract to develop minimum eligibility criteria and procedures. The contract which began in July, 1982, continued through June, 1983.

Purpose of the project. The purpose of the project was to develop and empirically validate eligibility criteria which can be used to identify community college students in need of learning disabilities services. The recommended eligibility criteria were to be based on the new definition of learning disabilities proposed for Title V for the California community colleges (Appendix A). The goals of the project were to: (a) specify educationally relevant attributes that differentiate learning disabled adults from low-achieving adults; (b) develop appropriate discrepancy/selection model(s) to certify the presence of LD in community college adults; (c) establish the magnitude of discrepancy and/or cutoff levels necessary to verify a learning disability among community college students; (d) complete a pilot study of possible eligibility criteria. This last goal differs from the original plan. In September, 1982, the Chancellor's Office recognized that it was impossible to complete a validation of the eligibility criteria during the first year, given the limited resources. Consequently, the project goal shifted to providing a field test of the eligibility criteria in a limited number of settings and proposing a plan for a future normative and validation study (see Recommendations section).

The proposed final product was to consist of step-by-step procedures to guide the identification process from the initial referral to the final identification decision. These procedures would operationalize the minimum eligibility criteria and utilize the decision models developed throughout the study. Once initiated, the procedures would result in increased statewide LD identification consistency and ensure a more equitable delivery of LD program services and allocation of excess cost funds to LD students in the California community college system.

Problems addressed. The development of minimum eligibility standards for identifying LD adults is surrounded by an array of complex issues. Unless each of these is dealt with systematically, validated identification procedures for LD populations are not likely to be generated. The major problems that have been addressed as part of the present project will be delineated in the following paragraphs.

1. Differentiating attributes have not been specified. At the heart of the LD identification issue is the ability to specify those learner attributes or characteristics that best differentiate the person with learning disabilities from others who are doing poorly in academic settings because of general low achievement. To date, researchers have been unable to identify any single attribute that is sufficient (by itself) to indicate a learning disability, or to specify a learning disability which is so debilitating that it warrants specialized educational opportunities. In short, existing attributes are, at best, equivocal indicators of a specific learning disability. This problem is exacerbated by the paucity of empirical data on the attributes which reliably differentiate the learning disabled individuals from the larger group of other low achievers. In spite of these circumstances the LD specialist is expected to decide whether or not a student is learning disabled. Not surprisingly, therefore, inconsistent identification practices and numerous selection errors have resulted.

2. Inappropriate assessment instruments are used in the identification process. Accurate identification of an LD population requires the use of appropriate psychometric testing instruments to quantify the decision-making process. A recent study sponsored by the Chancellor's Office (Ostertag & Baker, 1982; 1983) reported on the psychometric instruments

most frequently used for identification and assessment purposes. This investigation revealed two basic problems in current identification practices. First, several test instruments used suffer from insufficient technical adequacy (e.g., lack of age-appropriate norms) for a community-college population. Consequently, the application of such instruments will not yield valid data. Second, Baker and Ostertag (1982; 1983) reported great inconsistency in the use of selected instruments. That is, some colleges use a given instrument for identification purposes, whereas other colleges use the same test for making diagnostic inferences (e.g., specifying particular deficits in a student's math functioning).

3. Sufficiently powerful (sophisticated) model(s) for measuring achievement discrepancies have not been specified. The statistical concept of a discrepancy between a student's actual achievement (as measured by a standardized achievement test) and his/her predicted achievement (as measured by a standardized aptitude test) is one of the most commonly cited characteristics of a specific learning disability. In general, LD students are seen as achieving below their predicted achievement level. Of the numerous problems related to the operationalization of the discrepancy concept, the following are among the most critical: (a) How should the actual discrepancy score be calculated (e.g., raw score method, residual score method, true score method, regression method)? and (b) How should the discrepancy score be interpreted (e.g., statistical significance, educational significance?)

While simple to calculate, the more popular, traditional methods of measuring and interpreting discrepancy scores are fraught with naive and inaccurate assumptions. In part, the reliance on the simpler methods may stem from a lack of understanding of the more technically adequate methods. However, any method for calculating a discrepancy score must demonstrate sufficient technical sophistication to evaluate the influence of such statistical measurement factors as measurement error, regression effects, and increased variability of achievement scores with increasing age.

The effects of the above factors must be evaluated through a study of various discrepancy methods to avoid such selection errors as exclusion of a disproportionately large group of students whose aptitude scores are below or above average. Unfortunately, no empirical evaluation of these alternative discrepancy methods has been undertaken.

4. Other criteria must be specified in addition to aptitude and achievement measures. While the discrepancy between a student's assessed achievement level and predicted achievement level constitutes a commonly agreed-upon criterion for specifying a learning disability, it is not sufficient by itself for determining the presence of a learning disability. Without careful consideration of other factors, an identification procedure responsive to the needs of the LD population cannot be developed. Thus, the following problems arise from relying solely upon discrepancy criteria. First, even after the discrepancy scores of a group of students have been calculated and determined to be significant and severe, it is still unknown which of those discrepancies are due to a specific learning disability, low motivation, cultural differences, or educational disadvantage. Consequently, additional criteria are needed to differentiate the learning disabled population from other low-achieving groups. Second, most psychometric instruments used in discrepancy models are inappropriate for populations consisting of Limited English Speaking (LES) or Non-English Speaking (NES) individuals because these tests lack non-English versions

and norms. Thus, the validity of the discrepancy criterion must be specifically evaluated for these subgroups. Third, other components of the existing LD definition have not been operationalized. For example, although the "exclusion clause" of the LD definition emphasizes that the academic discrepancy may not be due primarily to such causes as severe emotional disturbance, environmental or economic disadvantage, cultural/ language differences, etc., no empirically based examplars are available to indicate how these criteria should be operationalized. In summary, a major need exists to empirically validate a set of factors to be used in conjunction with the discrepancy factors in making LD identification decisions.

5. Eligibility scores for student admission to LD programs must be established. Prior to PL 94-142 and the recognition of specific learning disabilities as a handicapping condition, the LD population was composed of a small, clinically defined group. However, as the popularity of programs for the learning disabled grew in recent years, the number of students being referred for LD services has grown at an even more rapid pace. Because of limited resources, cutoff levels are now being established for determining who should be eligible for such services and who should not. In the past, clinical judgment constituted a major identification strategy for LD populations. However, many of these clinical decision-making practices failed to systematically consider the values of key stakeholder groups (e.g., college administrators, LD specialists, Department of Finance personnel) in deciding the minimum eligibility criteria. The issue of setting eligibility cutoff scores is as central to the identification issue as the specification of which test instruments to use, for example. Appropriate cutoff levels can be established by using decision-

making methodologies that explicitly consider the values, concerns, motives, and interests of the key stakeholders involved. In addition, as these procedures are based on both clinical and statistical criteria they can take into consideration the rates of correct and incorrect classifications and assess the positive and negative consequences of any decision regarding cutoff levels.

6. Empirically validated data are lacking on existing identification procedures. To date, no comprehensive studies have been conducted within the learning disability field to validate current identification practices. Until such practices are submitted to rigorous empirical validation, the field will be unable to justify decision-making policy or intervention practices. It is imperative, therefore, that any proposed eligibility criteria be thoroughly tested to determine their efficacy for field application. It should be noted that numerous studies have provided essentially an "armchair" analysis of one or two components of the eligibility criteria and procedures. Often, what was offered as promising, innovative identificaton procedures turned out to be far less, and actually resulted in more confusion by promising much while producing little of significant utility. Until the problems previously mentioned have been collectively resolved and validated in an empirical study, the proposed identification process and procedures are "strawmen", i.e., they represent best guesses, and the practical consequences of a given model remain unknown. The California community college system is in the unique position of being able to systematically gather the data necessary for validating its selection procedures and empirically determining the requirements and consequences of their implementation. If this opportunity is ignored, the related problems will increase in number, range, and severity.

Project Participants

The Consortium for the Study of Learning Disabilities in California Community Colleges was created under the auspices of the Chancellor's Office of the California community colleges and charged with the task of developing specific procedures for assessing and identifying LD students in California's community colleges. The Consortium was composed of representatives from the Chancellor's Office, the California Association for Post-Secondary Educators for the Disabled (CAPED)--Learning Disability Division--and the University of Kansas Institute for Research in Learning Disabilities. The Consortium Steering Committee representing the Core Staff was composed of individuals directly responsible for executing the project. Steering Committee members included Laurel Best, Director, Educational Services Center, Cypress College and Representative from CAPED--LD Division; Drs. Don Deshler, Gordon Alley, and Daryl Mellard, University of Kansas Institute for Research in Learning Disabilities; Robert Howard, Coordinator/Administrator of Handicapped Students' Programs and Services, Chancellor's Office; Dr. Bruce Ostertag, Assistant Professor of Special Education, California State University at Sacramento; Dr. Ronald Baker, Coordinator/Enabler of Handicapped Student Services at Mira Costa College. The above individuals were selected by the Chancellor's Office based on their demonstrated interest and expertise in the research related to identification of adult-aged LD individuals in community college settings.

While the initial research plan was developed jointly by members of the Steering Committee, the responsibility for initiating the activities necessary to accomplish each goal was delegated among Steering Committee members. The KU-IRLD staff assumed responsibility for researching, developing, and evaluating the specific techniques and instruments to be used for data collection relevant to the project goals. The California-based staff, in turn, established contacts for collecting the necessary data within the community colleges themselves.

Recognizing the complex nature of this undertaking, the Steering Committee took considerable care in organizing the personnel and management aspects of the project. The following two criteria were considered critical in meeting the Consortium's goals.

First, given the array of issues and decision points associated with the development of an LD identification procedure, it was necessary to systematically solicit a broad base of input from key individuals, professional organizations, and state agencies. The Management Chart in Appendix B depicts how each project goal was supported by a Task Force(s) to facilitate the research and decision-making process. (Appendix C provides a complete listing of the members of the various task forces who actively participated in the project.) In addition, a State Advisory Committee was established to provide input to the Steering Committee. The listing of the individuals serving on the Task Forces and the Advisory Committee reveals project participants' representativeness and expertise.

Second, the need for technical expertise was encountered throughout the project. From the beginning, the Steering Committee has been committed to accessing experts with national/international reputation to obtain input at key junctures in the research process. Names of the consultants who provided such expertise are listed in the Management Chart (see Appendix B). No activity conducted in the learning disability field to date has utilized expertise equal to that tapped for the present project. In addition to providing input that was on the "cutting edge," these experts brought another critical dimension to the efforts of the project -- objectivity. That is, none of the experts had a vested interest in the direction or outcome of the project. This was felt to be a critical consideration since adherance to a rigorous set of scientifically defensible research procedures was deemed essential.

In summary, the Steering Committee views as one of the greatest project strengths of the broad base of input received from key experts and practitioners throughout the research study. In addition to an effective and necessary system of checks and balances, a process has been developed whereby some of the best thinking in learning disabilities and related fields was tapped to insure successful goal attainment.

Overview of the
California Community College
Learning Disabilities Programs

Prior to this study, Ostertag and Baker (1982) completed a comprehensive survey of the community college LD programs in California. The results of this survey provide a context in which the current study can be better understood. Specifically, all 106 California community colleges received and responded to a questionnaire soliciting information about the characteristics of their LD programs, as well as identification procedures and assessment techniques currently employed. The following is a summary of the major findings on each of these three areas.

Program characteristics. All identified LD programs were designated as formal or informal, formal programs being those in which all informal services were offered in additin to a) availability of an LD specialist; b) employment of a standard, routine assessment procedure; c) availability of special classes. Using these criteria, 80 of the 106 community colleges reported "formal" LD programs. An additional 13 colleges offered some type of "informal" program. Only 13 colleges reported not serving LD students. Approximately 8000 students were receiving some type of formal or informal LD service in 88% of the California community colleges at the time of the study.

A wide range of services was found to be available to community college LD students. The most common of these included special class offerings, tutorial support, and academic, personal, and/or career counseling. Other LD student supports provided by the majority of the community colleges (listed in order of frequency ranking) were diagnostic assessment, registration assistance services, remedial support services, lecture reproduction and comprehension aids, reading and notetaking assistance, class schedule modifications, and course curriculum revisions.

Individual Educational Programs (IEPs) were maintained on over 98% of the students in the formal LD programs. 21% of the community college programs reported that multidisciplinary teams were regularly employed to develop the IEPs even though these teams are not required for the development of IEPs at the post-secondary level. Over 93% of the community college programs provided inservice training for other faculty and staff.

Reading was ranked as the number one problem for students served in LD programs with writing and spelling as the next major difficulty area. Math and oral communications were ranked fourth and fifth, respectively.

Identification procedures. The identification procedures described by the respondents showed a number of consistencies particularly in the source of referrals and in the use of intake interviews with referred students. Thus, 81% of the programs conducted some type of standardized assessment as a follow-up to referrals, and over half of the programs included informal testing in their assessment procedures. Placement in the LD programs was decided in informal team meetings by the LD specialists or the Coordinator/Enabler of Handicapped Services. (Multidisciplinary teams are not required for post-secondary LD identification and placement.)

Assessment tools. Although a wide range of tests are used to assess learning disabilities, three published standardized tests were reportedly used in over 70% of the formal LD programs. In rank order of usage, these were: (a) the Wide Range Achievement Test (WRAT); (b) the Peabody Individual Achievement Test (PIAT); and (c) the Peabody Picture

Vocabulary Test (PPVT). Four additional tests were used by at least half of the respondents: Detroit Test of Learning Aptitude, Woodcock-Johnson Psychoeducational Battery (WJPB), Wepman Auditory Stimulus Test, Wechsler Adult Intelligence Scale-Revised, and Keymath.

Conclusions. The findings of the Ostertag and Baker Survey (1982) demonstrate several parallels among the California community college LD programs. Most noteworthy are the fairly consistent referral sources, use of intake interviews with referred students, assessment instruments used and the nature of the program services available.

Despite the above consistencies, insufficient evidence was found of a clearly defined process for identifying LD students and specific eligibility criteria that can be applied with a high degree of consistency. In their recommendations for what an identification process should include, Ostertag and Baker noted that many of the California community colleges are already adhering to most of their suggestions. Consequently, these recommendations were taken into account in the final drafting of identification procedures developed as the end product of this research.

Overview of Goals and Expected Outcomes

Four goals, each with its own set of expected outcomes, were developed at the start of the project. The completion of each goal represented a critical component of the development of procedures for identifying LD students in California's community colleges. The specific objectives and activities which led to the accomplishment of each goal are discussed in detail in the following sections. However, the goals and expected outcomes are introduced here in outline form as an overview of the major components of the project.

Goal 1. To specify educationally relevant attributes and technically adequate instruments that differentiate learning disabled adults from other low-achieving adults.

Intended Outcomes:

1. To specify a list of attributes which best differentiate LD from low-achieving students in community colleges;
2. To specify a set of reliable and valid instruments that can be used to operationally measure each attribute;
3. To provide an empirical validation of the criteria included in the California LD definition.

Goal 2. To develop appropriate discrepancy/selection model(s) to certify the presence of LD in community college (CC) adults.

Intended Outcomes:

1. To specify a discrepancy/selection model is appropriate for CC settings and which can be practically used to verify LD among CC populations.

Goal 3. To determine the magnitude of discrepancy and/or cutoff levels necessary to verify a learning disability among CC students.

Intended Outcome:

1. To establish a level of student performance on which LD can be verified on a consistent and agreed-upon basis.

Goal 4. To empirically validate the verifying properties (attributes, models, and discrepancy or cutoff levels) of a learning disability as it exists among community college students.

Intended Outcomes:

1. To determine the effects of the application of the specified discrepancy/selection procedures.
2. To produce a revised and validated selection process.

GOAL 1

Goal Statement: To specify educationally relevant attributes and technically adequate instruments that differentiate learning disabled adults from other low-achieving adults.

Problems addressed by Goal 1: LD differentiation attributes have not been specified. The lack of information about those attributes which can reliably differentiate LD from other low-achieving (LA) students in community college settings represented a major problem which had to be addressed before LD eligibility criteria could be established. The community college LD population has not been previously researched. In addition, only limited research has been reported on the general LD adult population. While identification procedures and eligibility criteria are available for younger-aged LD populations, these were considered inappropriate for adults. Furthermore, the generalizability of characteristics reported on adolescent LD students to community college LD students is questionable, since the latter group of individuals who, on their own volition, seek out post-secondary education despite their learning difficulties, most likely represent a unique subgroup of the heterogeneous LD population.

In addition, it was recognized that the primary difficulty encountered by LD practitioners did not lie in differentiating LD from normally achieving students, but rather in distinguishing LD from other low-achieving students. Therefore, it was imperative that, prior to setting eligibility criteria for LD, attention be focused on identifying and empirically validating those attributes which best differentiate LD from other low-achieving community college students.

Inappropriate assessment instruments are used in the identification process. The formulation of Goal 1 also took into consideration the need to identify valid and reliable psychometric instruments and procedures for measuring those attributes found to differentiate LD from low-achieving students. Thus, the initial intent of Goal 1 was to identify appropriate assessment instruments and then match instruments with the validated characteristics. Assessment practices employed in California's community colleges had been criticized (Ostertag & Baker, 1982) for using technically inadequate and age-inappropriate tests. In order for the identification of community college LD students to become more reliable, the instruments upon which eligibility decisions are based must be technically adequate and be used appropriatly and consistently.

Intended outcomes. The lack of empirically validated differentiating attributes of LD combined with the need to identify appropriate assessment instruments represents the central issues which Goal 1 was designed to address. Hence Goal 1 was formulated and its activities were designed to achieve two major outcomes:

1. specification of a list of attributes that best differentiate LD from LA students;

2. specification of a set of reliable and valid instruments that can be used to measure those attributes which are found to best differentiate LD from LA students.

Each of these outcomes would yield useful information for formulating minimum eligibility standards for placing students in California's community college LD programs.

Major activities. Each of the two projected outcomes of Goal 1 can be viewed as subgoals which, together, would accomplish Goal 1 and consequently resolve some of the major problems which have previously prevented the development of minimum eligibility standards for LD. Several activities were designed to culminate in the attainment of each outcome. Although a timeline for Goal 1 would show concurrent activities related to each outcome, it is easier to understand how each outcome was accomplished by considering activities relevant to each outcome separately rather than viewing a chronological accounting of all the activities associated with Goal 1.

Activities relevant to Outcome 1:

1. Compilation of a list of adult LD attributes reported in the research literature.
2. Development and distribution of a survey instrument to elicit LD adult characteristics from LD specialists (Survey 1).
3. Compilation of a list of educationally relevant attributes of LD adults based on responses to Survey 1.
4. Development of Survey 2 based on attribute lists resulting from the literature search and Survey 1.
5. Collection of LD content experts' estimations of the frequency with which the attributes in Survey 2 are observed in LD and LA adult populations.

Activities relevant to Outcome 2:

1. Initial screening of assessment instruments currently used in California's community college LD program.
2. Evaluation of assessment instruments by psychometric assessment experts.

Method

Assumptions

The activities relevant to specifying those characteristics which best differentiate LD from LA adults were based on the following assumptions. First, no single source of information should be considered reliable or inclusive enough to form the sole basis for a list of attributes of LD adults. Instead, attributes must be collected from a number of clinical and empirical sources and selected on the basis of multiple citations by LD experts and/or empirical data. Second, the most informed sources of attribute information are those individuals who work with LD adults in instructional or research settings. These professionals include LD specialists actively involved in teaching LD adults in community college programs and LD content experts recognized as having extensive experience with LD adults in research or instructional settings. Third, the final attribute list should include only those attributes which can be observed and measured in some standard manner. Although, historically, the exact nature of learning disabilities has eluded concrete measurement attempts, the goal to develop standardized identification procedures and eligibility criteria required that the attributes selected for identifying LD adults be observable and measurable.

Activity 1: Compilation of a list of LD adult attributes reported in the research literature.

Literature search procedures. A thorough search was conducted of the published literature to locate specific attributes found to characterize LD adults. Two KU-IRLD research assistants completed a computer-assisted literature search in four research indexes: (a) Educational Resources

Index Clearinghouse (ERIC); (b) Exceptional Child Educational Resources (ECER); (c) National Center for Mental Health (NCMH); (d) American Psychological Abstracts (APA). Descriptors used to locate articles included (a) learning disabilities; (b) adults; (c) young adults; and (d) adolescents.

The computer-assisted search, which included publication dates from 1966 through August, 1982, yielded 973 citations and abstracts. Each of these abstracts was reviewed and only those found to be relevant to the characteristics of LD adults were retained. Approximately 200 relevant citations were found.

Identification of the relevant citations was followed by a manual search to locate these articles so that descriptions of LD adult characteristics could be abstracted. Two research assistants searched three University of Kansas libraries for the pertinent articles. Interlibrary loans for articles unavailable in any of the Kansas libraries were also utilized. In all, 63 different journals were searched for the targeted articles plus additional articles listed either in reference lists or elsewhere in the journals being examined.

Not all the articles located in the manual search contained information relevant to the task of developing a list of attributes of LD adults, Also, as information was abstracted from the articles, it quickly became apparent that some articles were more data-based than others. Thus, it was necessary to develop screening criteria to help identify those articles that contained information specific to the purpose of the intended outcome. These criteria were: (a) the referenced subject samples must have reported ages above 17 years--11 months; (b) empirical findings relevant to the targeted LD population must have been reported or reviewed in the article--characteristics based on what appeared to be conjecture or opinion were not included; (c) only characteristics from published manuscripts were included; dissertations were not used unless the dissertation research had been reported in a published research article. Use of these screening procedures resulted in the collection of 59 research manuscripts suitable for use in developing a list of attributes of LD adults. A reference list of these materials is included in Appendix D.

Each of the selected 59 articles was analyzed for specific information relevant to the characteristics of LD adults. To standardize the recording of characteristic formation the Literature Search Analysis Form (see Appendix E) was designed. These record sheets were used to summarize the following information as it was located within the research articles: (a) identification of information relevant to the citation; (b) description LD employed in the paper--both conceptual and operational; (c) descriptions of marker variables for the LD sample (this information was collected using the descriptive, substantive, and topical markers for research recommended by Keogh, Major-Kingsley, Omoro-Gordon & Reid, 1982); (d) descriptions of and labels for the attribute/characteristics reported in the paper; (e) prevalence of the characteristic; (f) classification of the type of data being reported, e.g., true experimental, comparative, correlational, and professional opinion.

Several problems arose in abstracting information from the articles. In relation to the task of collecting empirically validated attributes of LD adults, the most important problem concerned the nature of the attributes reported in the literature. Often it was difficult to determine precisely how the attributes could be described in observable and/or measurable terms (e.g., vocational adjustment problems, dyslexic, dysgraphic).

Other attributes were overly specific (e.g., mean scores on some informal measure) or too broad in scope (e.g., social problems) to be useful. As a further complication, a wide variety of terms was used to describe what appeared to be similar phenomena or characteristics (e.g., reading disabled, dyslexic, dysgraphic, writing problems). Precise definitions or descriptions of the meaning of these terms were not always included in the articles. Occasionally characteristics had to be inferred from subject performance on an informal measure or rating scale.

Another major concern was the incomplete reporting of the characteristics of the research populations termed LD. Thus, criteria for classifying a subject as LD were not adequately reported in many of the articles. For the majority of the studies reviewed, the Keogh marker variable analysis (Keogh, et al., 1982) was virtually useless. Thus, it was often difficult to ascertain whether a study claiming to have used LD subjects was indeed reporting findings about LD subjects defined in the same sense as in the California CC LD definition. Although the reviewers attempted to abstract only those attributes that were clearly data-based and representative of LD rather than LA populations, it was sometimes difficult to be certain that the populations included in the studies were indeed learning disabled.

To ensure that only data-based attributes were included, the following precautions were taken: (a) only articles which clearly were presenting empirical findings were selected; and (b) the authors' descriptions of the attribute were initially recorded verbatim so that the attributes would be original descriptions rather than paraphrases. Although potentially discriminating characteristics may have been deleted by the stipulation of such stringent criteria, such criteria were the best protection against including the myths often associated with learning disabilities.

Procedures for compiling attributes. The next major task relevant to Activity 1 consisted of compiling the attributes abstracted from the literature search into a set of specific characteristics. The resulting characteristics would later be combined with the characteristics collected in Survey 1 (described in Activity 2) to develop the survey instrument for Activity 4. The following procedures were employed to abstract and categorize the literature-search attributes recorded on the analysis sheets.

A research assistant reviewed the analysis sheets and listed each characteristic separately on paper along with a citation code number and the source of the characteristic (i.e., LD definition, data-based characteristic, opinion-based characteristic, source unknown). Whenever an attribute appeared to embody a concept identical to one already listed from another source, the second citation was recorded next to the initial entry. However, attributes which were similar but included additional information were listed separately. This system enabled us to be concise without losing critical characteristics. In addition, a running count on the citation frequency of each attribute was maintained. Attributes from the 63 articles were included in this listing.

The next task was to sort the list of characteristics into the categories in the proposed Title V LD definition adopted by the California community colleges. As the characteristics were categorized, it became apparent that a number of them belonged in categories not included in the LD definition; thus, additional categories were added. The final categories and subcategories which emerged were:

I. Cognition
    A. Intelligence
    B. Information Processing

II. Achievement

   A. Oral Language

   B. Reading

   C. Math

   D. Written

III. Achievement/Ability Discrepancy

IV. Instructional Factors

V. Adult Adjustment

   A. Vocational

   B. Emotional

   C. Behavior

Finally, three members of the project staff collapsed all literature-based characteristics into a set of nonredundant, observable characteristics using the following procedures.

Procedures for Collapsing Literature-based Characteristics

1. Choose one of the categories (e.g., cognition, reading, math)
2. Examine the list of characteristics in the particular for content similarities.
3. Select one of the characteristics as best representing the content similarity which you have observed, or provide your own concept of those characteristics. The preferred criterion was to select an observable or measurable concept/characteristic.
4. Write the chosen characteristic/concept on your page under the appropriate heading.
5. Beside the characteristic/concept, write the number of the articles which included that characteristic. For example, the characteristic, poor spelling, was mentioned in articles

5, 8, 12. Write these numbers--5, 8, 12--beside the characteristic on your paper.

6. Proceed through the list in this manner including all characteristics for all categories. Cross out items as you use them.
7. Repeat steps 1-6 for each category of characteristics until all categories are complete.

This procedure resulted in a list of 75 adult LD attributes as reported in the published research literature. This list completed the tasks associated with Activity 1. (The characteristics resulting from this clustering procedure are included in Appendix F)

Activity 2: Development and distribution of a survey instrument to elicit LD adult characteristics from LD specialists - Survey 1.

In addition to the characteristics reported in the LD literature, we were interested in identifying CCLD adults' attributes as perceived by those individuals who have an opportunity to observe these students in instructional situations--community college LD specialists. This input was considered important for increasing the ecological validity of the final set of attributes and as well as serving as an important check on the characteristics obtained from the literature search.

Procedures for selecting community college LD specialists. Twenty-six LD specialists in the California community colleges were asked by Laurel Best to complete a survey of LD students' characteristics (Survey 1). Participants were selected on the basis of their exemplary performance as LD specialists in the community college system. The search for the 26 LD specialists was hindered somewhat as the community colleges were on summer vacation at the time the contacts were being made.

Design procedures for Survey 1. Survey 1 was an open-ended questionnaire, based on the five major components included in the proposed Title V LD definition. The instrument was designed to elicit LD specialists' perceptions of LD students' characteristics in the following areas: (a) intellectual ability; (b) academic achievement--listening comprehension, oral expression, written expression, reading recognition, reading comprehension, math calculation, math reasoning; (c) ability/achievement discrepancy; (d) cause of disability; and (e) exclusionary components.

Participants answered the following three open-ended questions related to each of the five LD definition categories.

1. How does this component contribute to the identification of CCLD students?

2. List classroom and/or test characteristics that: (a) are related to this component; (b) are commonly found among LD community college students'.

3. From the specific characteristics you listed, select one or more that best distinguishes LD from LA community college students. (Low-achieving community college students were defined as: (a) persons who received one or more failing grades in a required course on their most recent semester grade report. (Failing grades in activity courses are excluded from consideration, e.g., chorus, band, physical education.); 2) persons whose general achievement is at or below the 25th percentile as determined on a composite score of a standardized group achievement test; 3) persons who have no history of having been provided special education or vocational rehabilatation services.

In addition, each respondent was asked to: (a) describe any additional characteristics that useful in differentiating LD from LA adults in the community college; (b) provide estimates of the percentage of LD and LA students in the community college; and (c) provide basic demographic information about themselves. (The instrument used for Survey 1 is included in Appendix G)

Survey 1 was mailed August 5, 1982 with directions to return the completed questionnaires within one week from its receipt. The last survey was received September 28, 1982. A total of 13 surveys were returned.

Activity 3: Compilation of a list of educationally-relevant attributes of LD adults based on responses to Survey 1.

Procedures. As the surveys were returned, a research assistant compiled a master list of the characteristics reported in response to question 2: "List classroom and/or test characteristics that: (a) are related to this component (of the LD definition); (b) are commonly found among LD community college students."

Since the responses were in a narrative format, a procedure for isolating each characteristic was developed. The following procedure was employed:

1. Each response statement was read and the characteristic was identified.

2. Each characteristic was individually listed on a master list along with its survey identification code number. Descriptions of characteristics on the Master List matched the wording used in the survey as closely as possible. When paraphrasing was necessary, usually due to overly-elaborated responses, it was noted on the master list.

3. Multiple citations of the same characteristics were recorded by listing the survey code number for each citation on the Master List next to each characteristic.

A total of 413 separate characteristics was compiled from the 13 returned surveys. In order to reduce the master list to a more manageable size, the 413 characteristics were collapsed into 85 characteristics using clustering and cross-tabulation procedures. Thus, three independent judges clustered characteristics on the basis of similar content using the following directions which are similar to those used for collapsing the literature search characteristics.

Procedures for clustering Survey 1 characteristics. The following six-step sequence was used in organizing survey responses into logical factors.

1. Choose one of the component categories (e.g., oral expression, math calculation, math reasoning, intellectual ability).
2. Examine the list of characteristics for content similarities.
3. Select one of the characteristics as best representing the content similarity for the cluster of characteristics you have observed or, if none of the listed characteristics is adequate for this purpose, provide your own description of content similarity within the cluster of characteristics.
4. Write the clustered characteristic on your clustering sheet.
5. Beside each clustered characteristic, check the box numbers which correspond to the survey identification code numbers of each characteristic included in the cluster characteristic. For example, the characteristic Poor Spelling was cited on 9 surveys coded 1, 2, 4, 7, 8, 9, 11, 12, 13; therefore, you check the boxes for each of these code numbers.

6. Proceed through each category list until you have clustered all characteristics.

The clustering of the 83 characteristics from Survey 1 completed the tasks associated with Activity 3. (A list of clustered Survey 1 characteristics is included in Appendix H)

Activity 4: Development of Survey 2 based on attribute lists resulting from the literature search and Survey 1.

Procedures. The ultimate goal behind collecting and compiling the research-based and LD teacher-perceived characteristics of LD described in the previous activities was to arrive at a set of behavioral characteristics which could be used to differentiate LD from LA community college students. The next step in the narrowing process consisted of combining the two lists of attributes into a single list, which was to form the basis for a second survey in which various groups of LD experts were asked to estimate the frequency of each characteristic relative to LD and LA adult populations.

Thus, Survey 2 was generated using the results of Survey 1 and the literature search. The procedures used to combine the two sets of characteristics helped minimize the replication of characteristics while at the same time guarding against the loss of potentially LD/LA-discriminating data. These procedures are included in Appendix I.

First, the Survey 1 attributes were further refined by listing the 83 characteristics from Survey 1 on index cards and then sorting them into categories of similar content. Three judges reviewed the characteristics within each category to see if any characteristics needed to be collapsed further or if wording changes were necessary in order to make the described characteristic more observable. Having completed this step, the judges selected those characteristics that had been cited by two or more of the Survey 1 respondents. This two-respondent rule for selecting characteristics was adopted as a means of increasing the representativeness of each characteristic included in Survey 2. It seemed reasonable to expect that characteristics which are so representative of LD, as opposed to LA, that they are mentioned by more than one of the independent LD teachers in response to open-ended questions about the characteristics of LD would be most useful in discriminating LD from LA. A total of 61 Survey 1 characteristics met this criterion.

The literature-search characteristics were then added to this base using a different procedure. That is, each literature-search characteristic was compared to items on the two-citation characteristic list. If the content of the literature-search characteristic was already included in one of the Survey 1 characteristics, the literature search/Survey 1 match was noted and a point was added to the two-citation frequency count indicating that the same characteristic was cited in the LD research literature. On the other hand, if the literature-search characteristic was not represented in the Survey 1 list, the article from which the characteristic was abstracted was re-checked to confirm that the characteristic was indeed data based. If this was the case, the characteristic was added to the list; otherwise the characteristic was dropped. Application of this procedure to each of the 61 characteristics from the literature-search attribute list yielded 116 different data-based LD specialist-confirmed behaviorally stated characteristics associated with learning disabilities in adult-aged populations (Appendix J). These 116 characteristics formed the basis for Survey 2.

The goal behind Survey 2 was to determine which of the 116 characteristics commonly associated with LD would be useful in differentiating LD from other LA community college students. Recognizing that none of the 116 characteristics would likely be considered representative of every learning disabled student and no low-achieving student, it was deemed necessary to solicit information about each characteristic in a manner that would not require respondents to assign characteristics in an either/or manner to LD or LA, but would allow them to report their perception of the prevalence of the characteristic in the total LD and LA populations.

Therefore, the response format selected for Survey 2 required respondents to provide an estimate of the percentage of LD or LA students (depending on the form of the survey being completed) who demonstrate a given characteristic. Such frequency estimates have the advantage of providing an exact data point to be used for calculating exact percentage means for each characteristic. These means, in turn, can later be used in Bayesian likelihood ratios to determine the likelihood that a student who presents a certain cluster of characteristics is LD.

Another advantage of collecting percentage estimates for each characteristic is that the most powerful LD/LA differentiating characteristics readily emerge in a comparison between the mean percentages of LD and LA students for each characteristic. Those characteristics which yielded higher means for LD and low means for LA would be most useful in the LD selection procedures being developed, whereas those characteristics which are highly associated with both LD and LA could be deleted due to poor differentiation power.

It soon became apparent that it would be unreasonable to ask LD experts to provide frequency estimates on all 116 characteristics for both LD and LA students. Therefore, in an effort not to overtax Survey 2 respondents or to risk receiving incidence figures based on a desire to get the task over with rather than on considered judgment, a way was sought to reduce the demands on each respondent without forfeiting any of the 116 characteristics. To this end, Dr. John Poggio from the University of Kansas, who was consulted for a solution, recommended a matrix sampling procedure as a way to collect data on all 116 characteristics without having to include 116 characteristics on each questionnaire.

Using matrix sampling a set of survey instruments are generated rather than a single survey. While all 116 characteristics could be used within the total set, only 25 of them would appear on any one questionnaire. The 25 items included on a questionnaire are randomly selected and arranged on the questionnaires so that item appears a similar number of times within the whole set of questionnaires and each questionnaire contains a unique set and arrangement of items. This matrix sampling procedure was used to generate Survey 2. (Two complete sets of the surveys are included in Appendix K as examples.)

Since each respondent was to complete a survey on LD students (25 items) and another on LA students (25 items), "order effect" was feared. Order effect means that the second time a task if completed, it may be completed differently from the first time due to such factors as fatigue or familiarity. Therefore, if everyone had completed the LD survey first, for example, such factors might have influenced their responses on the second survey, the LA survey. To minimize the order effect, the surveys were arranged in an alternating pattern. For example, the first respondent was asked to respond to the LD survey first and then the LA survey. For the second respondent, the order was reversed with the LA survey appearing first followed by the LD survey. This alternating pattern of surveys, i.e., counterbalancing the order of administration, was followed for all respondents.

Activity 5: Collection of LD content experts' estimations of the frequency with which the attributes in Survey 2 are observed in LD and LA adult populations.

Procedures. Selection of respondents for Survey 2 was guided by a concern to be as inclusive as possible. As a result, Survey 2 was distributed to a wide variety of individuals, all of whom had experience with the LD/LA post-secondary population in various capacities. Care was taken to avoid duplication--only individuals who had not been included in another subject group were selected from subsequent group samples.

Each Survey 2 respondent filled out separate LD and LA surveys requesting prevalence estimates on 25 characteristics. The specific characteristics on the LD and LA surveys were randomly selected as a result of the matrix sampling procedure described previously. In addition to the two surveys, each respondent completed a Personal Data Sheet which elicited data on age, sex, professional role, and LD training plus information about the population with whom the respondent worked.

Eight groups of professionals were identified and contacted to serve as respondents to Survey 2. The first group was comprised of 39 LD content experts. Task Force members nominated these individuals on the basis of their reputation as teachers and researchers of LD adults. Each nominee was personally contacted by phone and asked to participate in the study. If the expert agreed to participate Survey 2 was mailed. (The complete list of LD content experts who responded is included in Appendix L).

The second group of respondents included 297 individuals listed as post-secondary LD program directors or teachers. This listing was obtained from an Association for Children with Learning Disabilities publication of nationwide post-secondary educational programs providing services to LD students. Each person listed in the directory was mailed a questionnaire; post-secondary LD programs from all 50 states were sampled. Out of 297 questionnaires mailed, 79 were returned in usable form.

The third group of respondents included all persons attending the LD division meeting at the October 1982 California Association of Post-Secondary Educators of the Disabled (CAPED) conference. All 36 individuals in attendance at the meeting completed and returned the questionnaires.

The fourth participant group was comprised of CAPED members. CAPED is a group of program administrators and instructors of handicapped students in the California community college system. All CAPED members listed in the 1982-83 membership list (who had not completed questionnaires at the CAPED conference in October) were mailed questionnaires. Of 96 surveys mailed, a total of 27 was returned.

A fifth group of respondents included the registered attendants at the 1981-82 California Community College Tutorial Association Conference. Out of 76 surveys distributed at the conference, 12 were returned.

The sixth group of respondents was selected from the Learning Resources Association of the California Community Colleges (LARC) 1981-82 staff directory. Seven organizations were represented in this directory: CCC Learning Resources Cooperative, California Media and Library Educators' Association, California Library Association--Community College Chapter, Community College Media Associations, Northern California Community College Computer Consortium, Western Educational Society for Telecommunications, and California Community College Chancellor's Office (Ex-Officia Member). The listings of resource staff in the directory were scanned and surveys were mailed to all individuals who were cited as being certified in education and serving as instructional staff in a resource program, and who had not previously received a survey as a result of being included in another group. One hundred fifteen surveys were mailed to selected individuals. Thirty four were returned.

The seventh group of potential respondents were identified from the directory of California Community Colleges listing individuals having primary administrative or instructional responsibilities for delivering LD services. One hundred twenty-two individuals identified in this manner received a survey. A total of 59 of 122 questionnaires were returned completed.

The last group of participants was comprised of individuals identified by a member of the California Department of Finance as being able to evaluate the characteristics of LA adults by virtue of their instructional experience within the California Community College system. Out of those listed, all but four individuals had already been sent surveys as a result of their membership in another group previously selected to participate in Survey 2. Two of these surveys were returned.

A total of 791 surveys were distributed either by mail or in person starting in October, 1982 and ending in January, 1983.

Data analysis. In January, 1983, 337 surveys had been returned--a return rate of 43%. Of the returned surveys, 22 were stamped "return to sender" by the postal service, while another 34 were returned unanswered or incorrectly completed. Data from 281 respondents were entered in the University of Kansas Honeywell 66/60 computer for subsequent analysis. Each data point was verified and corrected comparing the computer data file with the actual data on the survey.

At this point in the data analysis an error was discovered in the preparation of the surveys. As previously described, a computerized matrix sampling procedure was followed in generating the survey instruments. The detected error consisted of the repetition of a survey item on the computer file containing the survey items. As a result, 115 unique items were used in developing the surveys, rather than the expected 116. For the purposes of data analysis, the duplicated item was eliminated. Therefore, statistical analyses were completed on 115 items.

The analyses were considered as four separate steps. The first step, intended to evaluate for order effects, concerned the research question: since the order of administration of the LD and LA surveys was counterbalanced, was there a consistent difference between answering the LD survey first or second? Likewise, was there a reliable difference between responding to the LA survey first or second? This question was evaluated with Wilks' (lambda) statistic and the accompanying F-test.

The second step of the analysis consisted of evaluating for the possibility of another type of order effect. The corresponding research question was: were the responses different for the first items on the survey (e.g., items 1-5) than for the last items (e.g., items 20-25) regardless of item content? This question was based on a concern that participants might be responding to the items in some ordered pattern rather than to the uniqueness of the item content. For example, one ordered pattern might be that the percentages increased as a function of the item on the survey. Another possible pattern might be a tendency for respondents to give gradually increasing ratings to the items as they progressed from item one to item 25. To evaluate this issue, the items on the surveys were grouped into five clusters of five items each. The mean values for the five clusters were evaluated with a repeated-measures analysis of variance procedure. This procedure was applied to both the LD and LA surveys.

The third step of the analysis was to evaluate the 115 items for reliable differences between judgments concerning LD and LA students. Univariate t-tests were calculated on the percentage values given by the respondents to determine if the ratings for LD students were different from those given the LA students.

The fourth step of the analysis was a post hoc procedure. Since the survey respondents represented a wide range of backgrounds and experiences with learning disabled and other low-achieving young adults, the experimenters were concerned that these experiential differences would systematically influence the results of the item analysis. To evaluate this possibility the respondents were grouped according to the data they provided on the personal data sheet related to their contacts with the populations under study. More specificially, each respondent was requested to indicate his/her "primary role relative to the LD or low-achieving student" using the following four response options: community college LD specialist, college/university faculty, community college learning center teacher, and other. Since 84 of the 281 respondents indicated "other," an additional category was formed while three others were changed to better describe the groups of respondents. The resulting five groups were defined to include the following categorical roles: (a) LD specialists/teachers; (b) college/university faculty; (c) Resource teacher/clinician/non-LD instructor; (d) others; and (e) Administrator/director/coordinator. A one-way analysis of variance was completed on each of the 115 survey items to evaluate response differences by the five categorized groups.

RESULTS

The first analysis evaluated whether or not an order effect was present due to responding to the LD survey first or second and the LA survey first or second. Wilks' Λ (lambda) statistic was calculated and evaluated for significance with an *F*-test. For the order effect on the LD surveys, Wilks' Λ was calculated as 0.88666. The approximated *F*-value was calculated as 0.76187 with 25 and 149 degrees of freedom. This *F*value had a corresponding probability value of greater than .05. Therefore, it was judged not to be statistically significant. Due to missing data, 175 responses were used in this analysis rather than all 281 cases who responded. It was concluded that the respondents were not influenced in their responses by whether they completed the LD survey first or second.

A possible order effect on the LA surveys was evaluated similarly. The Wilks' Λ value was calculated to be 0.80005 with an approximate *F*-value of 1.36955. With 25 and 137 degrees of freedom, the probability value was greater than .05 and, therefore, judged not to be statistically significant. Due to missing values from the respondents, 163 cases were used in this analysis rather than all 281 cases who responded. It was concluded that there was no difference in the respondents' percentage estimates whether they responded to the LA survey first or second.

Another order effect was evaluated in the second analysis designed to determine whether or not the respondents replied differently to items as they progressed through the 25 items on each LD and LA survey. One-way analysis of variance procedures were applied to the LD survey responses and then to the LA survey responses.

The results of the analysis of the LD survey responses yielded an F-value of 0.52 with a corresponding probability value of 0.7176, with 4 and 460 degrees of freedom. Therefore, one might conclude that there were no reliable indications of a response pattern to the 25 items on the LD survey. In other words, the responses were given dependent on the item content and not the order of appearance.

A similar analysis was completed for the LA surveys. After applying the same analysis-of-variance procedures, the resulting F-value was 0.43. With 4 and 460 degrees of freedom, the probability value was determined to be 0.7877 and thus not statistically significant. The results were interpreted as indicating the absence of an order effect in the LA survey responses.

In the third step of the data analyses, the 115 individual survey items were evaluated for statistical significance. Based on the data analysis, 55 of 115 characteristics surveyed were statistically significant at the .10 probability level; on these 55 items, those respondents who completed surveys indicated that the LD population differed from the other low achievers. These items and the corresponding descriptive statistics are listed in Appendix M. Items which were found not to be statistically significant are listed in Appendix N.

The fourth step of the analysis determined if the five groups of respondents differed among themselves on their responses to the 55 items. Since different groups of professionals responded to the surveys, the question was whether or not any group or groups differed in their responses to the items. The results of this analysis revealed group mean differences on seven of the 55 significant items. Those items and the responses by group are listed in Table 1.

An additional analysis described in Appendix O was to logically group the 55 statistically significant items into associated factors. The results of this analysis is found in Appendix P. The 11 logically derived factors included: social, attribution, math, written communication, handwriting, complex reading skills, basic reading skills, variability, classroom performance, causal links, and oral communication.

Summary of Outcome 1. A comprehensive literature search and survey of LD specialists yielded a pool of 1156 attributes for distinguishing learning disabled adults from other low achievers. Those attributes were developed into a set of 115 survey items which were distributed to a wide range of professionals knowledgeable of post-secondary students. The 281 respondents completed the survey indicating the frequency with which a given attribute was characteristic of (a) learning disabled adults or (b) other low achievers. The data analysis yielded 55 items/attributes on which the respondents judged that LD and other low achievers differed. These items might be refined and later validated as a screening instrument for use by community colleges.

TABLE 1

Items with Differential Group Responses

Item 37: Has difficulty in correcting errors in writing.

	LD spec.	Faculty	Reg. Teacher	Others	Admin.
X̄	79.43	69.00	55.83	62.47	69.44
SD	16.5	22.1	24.85	19.31	26.45
SEM	2.36	4.94	7.17	4.43	6.23
N	49	20	12	19	18

Item 38: Has difficulty in organizing and sequencing written ideas.

X̄	76.90	70.82	72.90	69.91	55.36
SD	19.06	23.53	22.81	19.16	28.65
SEM	2.75	5.01	5.23	4,18	7.66
N	48	22	19	21	14

Item 64: Has average to above average intellectual ability.

X̄	81.16	74.47	65.22	65.39	59.17
SD	19.58	27.94	28.10	27.19	27.40
SEM	3.22	7.21	6,62	7.54	6.46
N	37	15	18	13	18

Item 66: Has variability in day-to-day performance.

X̄	57.59	56.05	70.26	54,20	37.33
SD	30.29	31.72	23.30	28.05	27.77
SEM	4.73	7.09	5.35	7.24	7.17
N	41	20	19	15	15

Item 94: Asks others to go some place with him/her.

X̄	47.93	50.40	45.00	51.38	28,94
SD	21.28	26.19	26.46	24.30	24,99
SEM	3.96	5.86	10.8	6.07	5.89
N	29	20	6	16	18

Item 99: Expresses few goals for the future.

X̄	16.92	14.07	33.75	10.00	66.67
SD	11.64	15.80	31.46	0.0	35.11
SEM	3.23	4.22	15.73	0,0	20.28
N	13	14	4	2	3

Item 103: Does not try hard to improve when criticized.

X̄	41.[illegible]	52.65	39.50	29.56	34.09
SD	23.80	28.47	19.90	18.13	23.96
SEM	3.63	6.37	4,97	4.27	7.22
N	43	20	16	18	11

Intended Outcome 2: Specify a set of reliable and valid instruments that can be used to measure relevant attributes.

Background. The second major outcome relevant to accomplishing Goal 1 was the specification of a set of reliable and valid assessment instruments which could be used to identify learning disabilities in community college-aged students. Although the initial plan had been to identify tests which could be used to measure the adult LD characteristics determined through Survey 2, it became apparent that many of the Survey 2 characteristics were not amenable to measurement using standardized assessment instruments. As a result, the plan shifted from test/characteristic matching to identifying tests that could be used to implement the discrepancy model of LD identification described in Goal 2.

The proposed Title V California definition of LD focuses on "Significant/severe discrepancy between the current level of developed intellectual abilities and academic performance." If ability/achievement discrepancy is to be used to identity LD, technically adequate and age-appropriate instruments must be used to assess a student's intellectual ability and academic achievement. Similarly, if a discrepancy model is to be used in determining student eligibility for LD services in the California community colleges, a set of reliable and valid assessment instruments must be specified for use in the discrepancy formula that is ultimately selected (Goal 2 provides a thorough discussion of the discrepancy concept and the recommended discrepancy formula). In determining which instruments to include, the KU-IRLD research team chose to rely on the judgment of individuals who are recognized experts on psychometric assessment. The following sections describe the resultant recommendations and how they were elicited from these experts.

Activity 1: Initial screening of assessment instruments currently used in California's community college learning disability programs

Procedures. Although the assessment experts were to make the final recommendations on which tests were the most appropriate for the purposes of discrepancy, the large number of potentially useful tests necessitated an initial screening to reduce the number of tests to be reviewed. The goal was to provide the experts with a list of tests which had already met certain basic requirements so that they could focus on identifying the best of such a list. The criteria discussed below were applied to the pool of potential instruments in order to (a) obtain appropriate measures of the traits under scrutiny; (b) eliminate as many sources of measurement error as possible using the instrument development procedures described in the test manuals; and (c) specify a set of technically adequate instruments

Screening Criteria:

1. Reliability coefficients (e.g., test-retest or split-half) of .85 or above must have been reported;

2. item content must have been judged appropriate to the purposes of the test;

3. concurrent/criterion validity coefficients must be available to provide a measure of the relationship between aptitude and achievement instruments r_{xy};

4. obtained scores must be in a form suitable for conversion to standard scores (X = 100, SD = 15);

5. a representative normative sample, recently obtained, was preferred for relevant comparisons.

These five selection criteria were applied to the pool of instruments generated using (a) the results of a survey study in California listing the most frequently used tests and assessment strategies employed in California community colleges (Ostertag & Baker, 1982; 1983); (b) recommendations of the IRLD staff; and (c) an agreement to include any new instruments which may not have been available for consideration in the Ostertag and Baker study. Application of the criteria yielded a subgroup of the following nine instruments:

1. Keymath Diagnostic Arithmetic Test
2. Peabody Picture Vocabulary Test--Revised
3. Peabody Individual Achievement Test
4. Raven's Progressive Matrices
5. Wechsler Adult Intelligence Scale
6. Wechsler Adult Intelligence Scale-Revised
7. Woodcock-Johnson Psychoeducational Battery--Parts I and II
8. Woodcock Reading Mastery Test
9. Wide Range Achievement Test

A tenth instrument, The Detroit Test of Learning Aptitude, was added because it was cited as a frequently used test in the identification of LD in the California community colleges.

The above list of 10 instruments is not extensive in terms of the number of tests for any one academic area (e.g., written expression). Also, tests are not included for each of the seven deficit areas in which a specific learning disability can occur. For example, oral expression and listening comprehension stand out as being unassessable based on the above tests. This paucity of appropriate instruments that meet the five screening criteria represents a major stumbling block in the development of discrepancy-based identification procedures and eligibility criteria for LD adults.

Activity 2: Evaluation of assessment instruments by psychometric assessment experts.

Procedures. The next step in the process of identifying appropriate ability and achievement assessment instruments was to ask recognized psychometric assessment experts to review the nine instruments which passed the initial screening plus the *Detroit*. Three assessment experts were asked to evaluate each of the nine instruments. The assessment experts were: (a) John Salvia--Pennsylvania State University; (b) Jerome Sattler--San Diego State University; (c) Dale Scannell--University of Kansas. The Mean Evaluation (Hoepfner et al., 1972) which provides a comprehensive method for evaluating test instruments with respect to measurement validity, examinee appropriateness, administrative usability, and normed technical excellence was selected as the format for these evaluations. (A copy of the Mean Evaluation Questionnaire is included in Appendix Q). Mean Evaluation Questionnaires were mailed on January 28, 1983.

RESULTS

The following sections contain summaries of the evaluations provided by the assessment experts for each of the 10 instruments. The tests have been classified as either aptitude or achievement measures to facilitate comparisons among tests of similar uses. Results are reported in terms of the five screening criteria (reliability, item content, r_{xy} aptitude/achievement correlations, standard score conversions, and norms). (Actual evaluations and correspondence from the assessment experts are included in Appendix R).

I. Aptitude Instruments:

A. Wechsler Adult Intelligence Scale (WAIS)

(1) Reliability:

Internal consistency reliability was reported at .97 for full-scale IQs, and .93-.94 and .96 for verbal and performance IQs, respectively.

(2) Item content:

Appropriate according to measurement experts. According to Satter (1982), "It is the best test for the assessment of community college age individuals" (p. ?).

(3) r_{xy}:

A measure of the relationship between the WAIS and the Woodcock-Johnson Psychoeducational Battery (WJPB), (Tests of Academic Achievement), was reported in the technical manual of the WJPB (see the relevant section of this report).

(4) Scores:

Scores are converted to standard scores with a mean of 100 and a standard deviation of 15.

(5) Norms:

Norms were obtained by employing stratified samples of groups considered representative of the adult populations of the United States at the time of the 1950 census.

Summary

The Wechsler Adult Intelligence Scale is an intelligence scale which includes both verbal and performance aptitude subscales. It was designed for use with individuals over 16 years of age. Sattler noted that the performance subscale may be used for individuals with limited English language proficiency. Original norms appear to be excellent although somewhat outdated and in need of revision. Overall, however the WAIS has acceptable technical adequacy and is appropriate for assessing the scholastic aptitude of community college-aged populations.

B, Wechsler Adult Intelligence Scale - Revised

(1) Reliability:

Test retest reliabilities were reported for Verbal, Performance, and Full-Scale IQs in two age groups: 25-34 (V = .94, P = .89, FS = .95) and 45-54 (V = .97, P = .90, FS = .96).

(2) Item content:

Approximately 80% of the items in the revised WAIS were retained from the original standardized version of 1955 and are considered appropriate to assess general intelligence.

(3) $\underline{r_{xy}}$:

No measure could be found of the relationship between the WAIS-R and an achievement instrument.

(4) Scores:

Sums of scaled scores are converted to Full-Scale and Verbal and Performance IQs with a mean of 100 and a standard deviation of 15 for all age groups.

(5) Norms:

The standardization was based on a representative sample of adults (16 years 0 months through 74 years 11 months) using a stratified sampling plan. Stratification variables included age, sex, race, geographic region, occupation, education, and urban-rural residence.

Summary

The Wechsler Adult Intelligence Scale-Revised is an intelligence scale which yields verbal and performance subscales and is appropriate for use with persons aged 16 to 75. Studies are needed to provide an index of the relationship between this instrument and suitable achievement tests (r_{xy}).

I.C. Detroit Tests of Learning Aptitude:

(1) Reliability:

Total test-retest reliability is reported in the manual as being .96 after 5 months with a sample of 48 pupils.

(2) Item content:

Item content (specific to 19 subtests) was considered by the authors to be appropriate for ages 3 years through adulthood; however, no method of determining appropriateness was described.

(3) r_{xy}:

The relationship (T_{xy}) between the Detroit and tests of achievement is undetermined.

(4) Scores:

Raw scores are converted into Mental Age scores. No standard error of measurement was reported. Neither standard scores nor percentile ranks by age were available.

(5) Norms:

The initial standardization sample included 50 pupils at each age level (levels unspecified). Representativeness of the sample was claimed by the authors, but no selection procedure was described. Standardization took place in the Detroit public schools.

Summary

The Detroit Tests of Learning Aptitude is a norm-referrenced, standardized test containing 19 subtests designed to provide information about several specific learning abilities. The authors consider the test to be appropriate for ages 3 years through adulthood. While test-retest reliability appears adequate, the instrument is unsuitable for determining severe discrepancy for the following reasons:

(a) Content appropriateness has not been determined; (b) no studies are reported to provide an indication of the relationship between this test and tests of academic achievement; (c) no standard scores or percentile ranks by age and no standard error is presented; and (d) no information is available to indicate whether the norming sample was representative of the overall United States population.

I.D. Peabody Picture Vocabulary Test - Revised

(1) Reliability:

Test-retest reliability coefficients for single age groups' standard scores ranged from .54 to .90. Median reliability was .82 for form L with the adult sample.

(2) Item content:

In the revised test Sattler (1982) maintained that items were selected to provide a balance of "nouns, gerunds, and modifiers in 19 content areas." Words which showed effects by sex or race were eliminated or counterbalanced.

(3) r_{xy}:

measure of relationship between the PPVT-R and a test of aptitude or achievement was provided.

(4) Scores:

Raw scores may be transformed into standard scores with a mean of 100 and a standard deviation of 15.

(5) Norms:

The standardization sample was nationally representative and included 4,200 students, ages 2½ through 18 and 828 adults. The student sample was selected on the basis of geographical region, occupation, sex, race, and community size. The adult sample approximated census data (1970).

Summary

The Peabody Picture Vocabulary Test - Revised (PPVT-R) is an individually administered, standardized test of receptive vocabulary. According to its authors the test can be used to measure achievement (English vocabulary acquisition and scholastic aptitude (verbal ability). The instrument does not represent a test of general ability. While it appears to meet the five criteria for technical adequacy in its aptitude function for adults, the PPVT-R should not be used as the only aptitude test in determining severe discrepancy because of its limited item content. Also its use is counter-indicated for limited-English speaking individuals.

I.E. Standard Progressive Matrices

(1) Reliability:

Test-retest reliabilities range from .83 to .93 across age groups as follows: under 30 years - .93; 30 to 39 years - .88; 40 to 49 years - .87; and 50 and over - .83.

(2) Item content:

Item content consists of 60 figures (5 sets of 12) considered appropriate for children (aged 6 and older) and adults to elicit reasoning with figural material. The content is not deemed appropriate for measurement of general intelligence. No measure of relationship between the Standard Progressive Matrices and an appropriate test of achievement could be found.

(4) Scores:

Raw scores are converted to percentiles.

(5) Norms:

Standardization, reported in 1960, used representative samples of British populations ranging in age form 6 years to 65 years. No standardization has been conducted with populations in the United States.

Summary

The Progressive Matrices test provides a measure of nonverbal reasoning. It is a standardized test consisting of 60 figural patterns which must be completed by employing comparisons, analogical reasoning, spatial perception, permutation and other relations. It is not considered a test of general intelligence. The Progressive Matrices have not been standarized in the United States, have dated norms for a British sample, and do not provide a measure of r_{xy} with an appropriate achievement of test. For these reasons, the instrument is inappropriate for assessing severe discrepancy in community colleges populations in the United States.

I.F. Woodcock-Johnson Psychoeducational Battery (W-SPB) (Cognitive Ability - Part I):

(1) Reliability:

Internal consistency reliability (ages 20-39 years) was .98 for the Full-Scale aptitude.

(2) Item content:

Specific information regarding the selection of item content for the 12 subtests was unavailable in the author's manual. Items were selected for their high correlations with achievement measures that are not confounded by the tests.

(3) r_{xy}:

The W-JPB technical manual contains information about the relationship between the W-JPB - Part I (Aptitude) and the W-JPB - Part II (Achievement). r_{xy} is reported.

for each achievement cluster (age: 20-39) as follows: Reading = .78; Mathematics = .72; and Written language = .72. Relationships of the W-JPB Part I (Full Scale) to other achievement tests was reported by subtest areas: PIAT Reading = .78; Mathematics = .79; and Written Language = .59, WRAT Reading = .74; Mathematics = .71; and Written Language = .71.

(4) Scores:

Cluster scores obtained from combined subtest performances constitute the recommended unit of analysis for interpretation. Cluster scores can be converted to percentile ranks or standard scores for comparisons.

(5) Norms:

All parts of the W-J PB used the same representative sample drawn from 49 communities. Of the total sample of 4,732 subjects, 3,900 comprised the school sample whenever the remaining 832 were composed of preschool and adult ages. The following stratification variables were used: sex, race, occupation, geographic region, and type of community (urban-rural).

II. Achievement Instruments

II.A. Woodcock-Johnson Psycho-Educational Battery: (Achievement Battery - Part II):

(1) Reliability

Internal consistency reliabilities were reported for cluster scores (age 20-39): Reading = .97; Mathematics = .94; Written Language = .97.

(2) Item content:

Items were developed with the help of curriculum consultants, experienced teachers, and other experts. The authors stated that 10 subtests were designed to cover the major areas of scholastic achievement.

(3) r_{xy}:

tionships between the WJPB Part I (Aptitude) and Part II (Achievement) are reported for achievement clusters: (ages: 20-39): Reading = .78; Mathematics = .72; and Written Language = .73. The degree of relationship between the WJPB Part II and the WAIS was also reported for achievement clusters: Reading = .66; Mathematics = .75; and Written Language = .65;

(4) Scores:

Percentiles and standard scores may be derived from cluster scores (scores from combined subtest performances). Cluster scores are recommended as the most preferred unit of analysis.

(5) Norms:

All parts of the WJPB were normed on the same representative sample of 4,732 subjects in 49 communities as noted in the standardization discussion of Part I.

Summary

The Woodcock-Johnson Psychoeducational Battery can be used to assess both cognitive ability (Aptitude) and achievement. Application of the five selection criteria yielded satisfactory outcomes for Part II (Achievement). Part II can be used appropriately as the achievement instrument to determine severe discrepancy with community college-aged students. Part I, while appearing to satisfy the reliability requirement, is limited in the range of cognitive abilities it samples. This limitation should be considered when selecting an aptitude measure for use in discrepancy determination.

II.B. Peabody Individual Achievement Test (PIAT)

(1) Reliability:

Test-retest reliabilities are reported for grades K-12. At grade 12, reliability coefficients for the five subtests ranged from .63 to .86 with a total reliability of 92 and a median reliability of .79. Subtest reliabilities were as follows: Mathematics = .84; Reading Recognition = .86; Reading Comprehension = .63; Spelling = .75; and General Information = .73.

(2) Item Content:

Item content was derived from reviews of curriculum materials used at grade levels K-12.

(3) r_{xy}:

A measure of the relationship between the PIAT and the WAIS (full scale) was reported in the Woodcock-Johnson Psycho-educational Battery technican manual for: Mathematics(.71); Reading (.68); and Written Language (.54).

(4) Scores:

Four types of scores can be obtained for the total test and the five subtests including standard scores based on mean of 100 and a standard deviation of 15.

(5) Norms:

The standardization sample (2,899) was composed of at least 200 children at each of 13 grade levels (K-12) based on geographical location and community size. Blacks comprised 11.3 percent of the sample; occupations were representative of the population at the time of the 1967 census.

Summary

The Peabody Individual Achievement Test is an individually administered test which provides a broad range of information on achievement in five subtests: Mathematics, Reading Recognition, Reading Comprehension, Spelling, and General Information. Reliabilities of specific subtests are too low to meet the test selection criteria for technical adequacy. Also, the absence of individuals older than high-school seniors in the standardization sample represents a limitation on employing this test with community-college populations.

II.C. Wide Range Achievement Test (WRAT)

(1) Reliability:

Reliability data presented in the manual (1978) consist of split-half reliabilities for seven undefined groups. Reliability coefficients reach .94 or above for all groups across all three subtests:

Level I (ages 5-11): Reading = .98; Spelling = .96; and Arithmetic = .94.

Level II (ages 12-24): Reading = .98; Spelling = .97; Arithemtic = .94.

(2) Item content:

Item content samples a highly restricted range of behaviors across three subtests. Obtained raw scores are reported to increase with age. Validity coefficients were obtained between the WRAT and other achievement tests for grade 1 through 6.

(3) r_{xy}:

r_{xy}: A measure of the relationship between this instrument and the WAIS is reported in the 1978 manual for each subtest. It differs for males (N = 200) and females (N = 200) ages 18 to 24 years:

	WRAT Males			WRAT Females		
	Reading	Spelling	Arithmetic	Reading	Spelling	Arithmetic
WAIS IQ	.84	.82	.83	.87	.83	.83

(4) Scores:

Three types of scores are available for each subtest: grade equivalents, percentiles and standard scores, with a mean of 100 and a standard deviaton of 15.

(5) Norms:

According to the 1978 WRAT manual no attempt was made to provide for a nationally representative sample. The 1978 standardization used a sample of 50% males/50% females, which included ages 5 through 64. The exact size of the sample was not reported.

Summary:

The Wide Range Achievement Test is an individual, norm-referenced, achievement test which provides limited information on achievement in reading, spelling, and arithmetic. Reliability coefficients reach desirable levels and the instrument appears acceptable for assessing achievement for use in severe discrepancy determinations. The major exception to the instrument's suitability for with community-colleges populations arises from the authors' failure to adequately define the norm sample. Users of the instrument should be aware of this limitation.

II.D. Woodcock Reading Mastery Tests

(1) Reliability:

Internal consistency reliability was reported at .99 for total reading. Subtest reliabilities ranged from .02 to .97 at the 7.9 grade level.

(2) Item content:

Item content for certain subtests may not be appropriate for community college-age populations, e.g., students in the upper ages of the norming sample correctly answered nearly all the items on the letter identification subtest.

(3) r_{xy}:

measure of the relationship between the *Woodcock Reading Mastery Tests* and the WAIS (Full Scale) was reported in the technical manual for the *Woodcock-Johnson Psychoeducational Battery* to be .82.

(4) Scores:

Both mastery ratings and traditional standard scores are provided.

(5) Norms:

The norm sample was obtained over a two year period in 50 school districts in the United States. The sample appears to be representative of the population in terms of community size, race, years of schooling, income, and occupation. In a total sample size of 5,252, grades K-12 were represented.

Summary

The Woodcock Reading Mastery Tests are intended to assess skill development in five areas of reading achievement. While the standardization was found to be adequate for grades K-12, it did not include samples beyond the 12th grade. Together with less-than-desirable reliability and inappropriate item content for specific subtests this inadequacy makes the instrument questionable for use with the community-college population. Also, it should not be the only reading achievement measure used to identify a severe discrepancy for any individual.

II.E. Key Math Diagnostic Arithmetic Test

(1) Reliability:

Internal consistency reliabilities (K-7) ranged from .39 to .90 by grade and subtest. Total test reliability for 209 items reached .96. No information was reported for test-retest stability.

(2) Item content: Item content was carefully selected based on the authors' review of 10 major programs of mathematics and field-testing of the specific items with more than 3,000 children (grades K-8) including 1,400 mentally retarded children in 20 states. In a separate study items were sequenced according to difficulty with elementary and junior-high school students.

(3) r_{xy}:

measure of the relationship between the Key Math and an aptitude test could be found for adult ages. The following information was available for a younger sample (Grade 5) using the WISC as the predictor = .48.

(4) Scores:

Raw scores are converted to grade equivalents. Supplemental tables which supply percentile ranks are available.

(5) Norms: The standardization sample included 1,222 children grades K-7 in 42 schools representing 21 school districts across 8 states. Both schools and students were randomly selected. Information was not available regarding type of program enrolled in, socioeconomic status, parental occupation, or education.

Summary

The Key Math Diagnostic Arithmetic Test is an individually administered, standardized test designed to diagnose math skills at grade levels K-8. This instrument's suitability for adult populations has not been determined in relationship to any of the five criteria. Studies are needed to determine the test's usefulness for identifying learning disabled community-college populations.

Summary of Outcome 2. The intent of the preceding activities was to identify those standardized assessment instruments which were technically adequate for use in a statistical discrepancy method or in assessing characteristics associated with learning disabilities. The results suggest that: (a) few technically adequate instruments appropriate for the community college population are available; (b) the Wechsler scales and possibly the Woodcock-Johnson Psychoeducational Battery, Part I, are appropriate indicators of academic potential; (c) the Wide Range Achievement Test and the Woodcock-Johnson Psychoeducational Battery, Part II, are appropriate measures of academic achievement; (d) bias may be evident in using these tests with selected subpopulations, e.g., head trauma; and

(e) the tests should be cross-validated with the community-college population.

Summary of Goal 1:

The purpose of Goal 1 was to specify a set of attributes that could reliably distinguish between LD adults and other low achievers and determine which formal assessment instruments are technically adequate for the community-college population. A survey procedure was followed in the identification of 55 attributes on which the respondents indicated that LD adults could be distinguished for other low achievers. These items are considered tentative and in need of validation with actual student populations.

Three experts in educational assessment independently evaluated the technical adequacy of 10 commonly used formal assessment instruments. The experts' ratings indicated that most of the tests lack technical adequacy for use with community-college populations. Two aptitude measures and two achievement measures met the criteria for technical adequacy.

GOAL 2

Goal Statement: To develop appropriate discrepancy/selection model(s) for verifying the presence of LD in community college adults.

Problems addressed by Goal 2: Sufficiently powerful (sophisticated) model(s) for measuring aptitude-achievement discrepancies have not been specified. The statistical concept of a discrepancy between a student's actual achievement (as measured by a standardized achievement test) and his/her predicted achievement (as measured by a standardized aptitude test) is one of the most commonly cited characteristics of a specific learning disability. In general, LD students achieve below their predicted achievement level. Of the numerous problems related to the operationalization of the discrepancy concept, the following are among the most critical: (a) how should the discrepancy score be calculated (e.g., raw score method, residual score method, true score method, regression method)? and (b) how should the discrepancy score be interpreted (e.g., statistical significance, educational significance)?.

While simple to use, traditional methods of measuring and interpreting discrepancy scores, are fraught with numerous inaccurate assumptions. Any method for calculating a discrepancy score must demonstrate sufficient technical sophistication to evaluate the influence of such statistical measurement factors as measurement error, regression effects, and increased variability of achievement scores with increasing age. In part, the reliance on simpler methods may stem from a lack of understanding of more technically adequate methds.

The effects of measurement factors must be evaluated through a study of various discrepancy methods. Unless the methods are evaluated, numerous selection errors will occur, such as exclusion of a disproportionately large group of students whose aptitude scores are below or above average. Unfortunately, no empirical evaluation of these alternative discrepancy methods has been undertaken.

Other criteria must be specified in addition to aptitude and achievement measures. While the discrepancy between a student's assessed achievement level and predicted achievement level constitutes a commonly agreed-upon criterion for specifying a learning disability, it is not sufficient by itself to determine the presence of a learning disability. Unless other factors are carefully considered, the identification procedure will not be responsive to the needs of the LD population. The following problems arise from relying solely upon discrepancy criteria. First, even after the discrepancy scores of a group of students have been calculated and determined to be significant and severe, it is still unknown which of those discrepancies are due to a specific learning disability, low motivation, cultural differences, or educational disadvantage. Consequently, additional criteria are needed to differentiate the learning disabled population from other low-achieving groups. Second, most psychometric instruments used in discrepancy models are inappropriate for use with Limited English Speaking (LES) or Non-English Speaking (NES) individuals because these tests lack non-English versions and norms. Thus, the validity of the discrepancy criterion must be evaluated for these subgroups. Third, other components of the existing LD definition have not been operationalized. For example, although the "exclusion clause" emphasizes that the academic discrepancy may not be due primarily to such causes as severe emotional disturbance, environmental or economic disadvantage, cultural/language differences, etc., no empirically based examplars are available to indicate how these criteria should be operationalized. In summary, it is evident that a major need exists to empirically validate a set of factors for use in conjunction with the discrepancy factors in making LD identification decisions.

Criteria for determining eligibility must be integrated into a comprehensive model. If the two previous problems can be effectively solved, discrepancy model and other factors on which to judge a student's eligibility for LD services will be available. However, the practitioner is still confronted with the problem of determining how to weigh the eligibility factors. For example, to be eligible for LD services a student must meet all the criteria or is it sufficient to meet three of five criteria? Clearly, it is necessary to not only provide eligibility criteria, but to integrate those criteria into a set of selection procedures that appropriately evaluate all of the relevant information for determining a student's eligibility for LD services.

Summary of problems related to Goal 2. Three critical problems underlie the development of a selection model for determining a student's eligibility: (a) selecting a sophisticated aptitude-achievement discrepancy model, (b) developing a selection model that includes other factors appropriate for identifying the heterogeneous LD population, and (c) determining a qualitative method that evaluates and weighs the multiple criteria for determining a student's LD eligibility.

Intended outcome. This project goal focused on one outcome: to provide a discrepancy and selection model which is appropriate for community college settings and can be used to certify a learning disability in the community college population. Since the development of a discrepancy model and selection model were both major tasks involving multiple activities, the discussion will treat each model separately.

Major activities. The attainment of the intended outcome was accomplished through the following activities:

Discrepancy model activities:

1. Identification of existing discrepancy models
2. Evaluation of current discrepancy models and generation of alternative models
3. Selection of the discrepancy model judged most appropriate for the community college setting

Selection model activities:

1. Identification of factors that are potential additions and alternatives to the discrepancy model
2. Development of a research methodology with a design demonstrating sufficient statistical power and utility.

Discrepancy Model Methods

Assumptions

In the previous discussion of problems addressed by Goal 2, a number of statements about the discrepancy model followed from specific assumptions. While several procedures are available for calculating a discrepancy score, our suggestions are based on the assumption that any discrepancy method appropriate to the community colleges should consider sources of errors in the measurement that would lower the reliability of the discrepancy score and correct for such errors.

We also assume that regardless of the discrepancy method, the presence of a significant aptitude-achievement discrepancy score is not sufficient for determining a learning disability. The need for further information may perhaps best be illustrated through a metaphor from medical science. For example, when a person is administered a tuberculin skin test, the results are either positive or negative. If the results are negative, the patient is considered free from tuberculosis. Likewise, if a student receives a valid evaluation of aptitude and achievement, and the resulting discrepancy score is not significant, the student would not be considered learning disabled.

In the case of the tuberculin test, if the skin test is positive, the conclusion is not automatically drawn that the patient has tuberculosis. Further diagnostic exams, e.g., a chest x-ray, are required to make such a determination. Likewise, if a student demonstrates a significant discrepancy score, it does not necessarily indicate a learning disability. Further diagnostic information is required to ascertain whether the significant discrepancy reflects limited educational experiences, mental retardation, cultural differences, emotional differences, etc. Therefore, while a significant aptitude-achievement discrepancy is necessary for determining a learning disability for most students, it is only one of the criteria.

The last assumption underlying the use of statistical discrepancy methods is that the significance of a discrepancy score must be based on statistics, not solely on the clinical judgment of the evaluator. Traditionally, test performance, e.g., discrepancy scores, was evaluated by the clinician to determine whether or not it was of educational or practical significance. In a statistical discrepancy method, this decision is made using a frequency distribution, not the clinician's subjective judgment. The frequency distribution indicates the probability of any discrepancy score occurring. The student's individual discrepancy score is evaluated for significance by comparing it to this known distribution, and judging whether or not the discrepancy score occurs frequently enough to be of practical significance.

Activity 1: Identification of Existing Discrepancy Models

Participating consultant. The steering committee recognized the need for having a consultant with extensive skills and practice in the areas of educational measurement and multivariate statistics. Since Dr. John Poggio, Chairman of the Department of Educational Psychology and Research at the University of Kansas, had previously been involved in collaborative efforts with KU-IRLD, Dr. Poggio was asked to provide (Appendix S) consultative assistance to the project.

Procedures. Project staff recognized the diversity of methods that potentially could be useful in calculating discrepancy scores. Some of these methods have been discussed as reading expectancy formulas, learning quotients, achievement quotients, gain score formulas, etc. To increase the probability that information relevant to discrepancy was evaluated, three information-gathering procedures were followed: (a) conducting a literature search; (b) contacting state and federal agencies; and (c) contacting other researchers currently investigating discrepancy methods.

The literature review involved computer searches from two different bases (ERIC and Psycinfo searches). In addition to 65 literature citations obtained from these searches, at least nine others were from the reference lists of the articles retrieved through the computer search.

Copies of these articles were reviewed by KU-IRLD staff to gain a thorough understanding of the factors relevant for considering discrepancy methods as well as potential pitfalls of their application.

By contacting education agencies in several states (i.e., Iowa, Michigan, California), project staff learned how other groups were proceeding in operationalizing aptitude-achievement discrepancy methods for identifying learning disabled students. In general, the results of this information gathering indicated that (a) discrepancy methods were being adopted; (b) current discrepancy methods were more sophisticated than those previously applied, e.g., calculations assessed regression effects, errors of measurement, prediction error, etc.; and (c) the state agencies were exerting tighter controls on students identified as learning disabled.

The third information source, initiated by Dr. Gordon Alley of KU-IRLD, consisted of personal contacts with researchers currently investigating discrepancy methods. Such contacts provided an opportunity to exchange information and to query concerning specific issues. As part of this information gathering several individuals were contacted, e.g., Dr. Lee Cronbach at Stanford University, Dr. Gerald Hanna of Kansas State University, and Dr. Lonnie Wilson at the Iowa LD Identification Project.

Based on information gathered from the three major sources outlined above, KU-IRLD staff prepared a document entitled Questioning Current Statistical Discrepancy Models for Identifying Learning Disabilities (Appendix T). The document reviewed three alternative methods of calculating discrepancy scores and four sets of issues: (a) determining the educational significance of a discrepancy score, (b) determining the appropriate population for calculating the values needed in the discrepancy methods, (c) determining the appropriate measure of predicted achievement, and (d) determining the preferred statistical discrepancy method. The document was used in obtaining responses from the measurement experts involved in Activity 2.

Activity 2: Evaluation of Current Discrepancy Models and Generation of Alternative Models

Participating consultants. Three individuals were asked to evaluate the technical issues concerning alternative methods for calculating discrepancy scores. Dr. John Poggio of the University of Kansas recommended and contacted (Appendix S) potential consultants for Goal 2. As a result, the following individuals agreed to participate: Dr. Huynh Huynh at the University of South Carolina, Columbia, S. C.; Dr. Robert Linn at the University of Illinois at Urbana - Champaign, IL; and Dr. Nancy Petersen at Educational Testing Service (ETS) in Princeton, N. J. These individuals were sought because of their recognized expertise in educational measurement issues.

Procedures. Following an acknowledgment from Drs. Huynh, Linn and Petersen to participate, they were mailed a copy of the paper, Questioning Current Statistical Discrepancy Models for Identifying LD (Appendix T) with a request to respond to the materials.

Their responses provided specific information concerning the factors to be considered when applying any discrepancy method for determining program eligibility. The responses to the specific questions are found in Appendix V.

The experts' responses led the Steering Committee to query several other related issues through conference calls and correspondence with the consultants. Since these extended discussions treated technical issues of discrepancy methods, a treatment of them here would not be useful for most readers. Instead, the sequence of letters between KU-IRLD staff and Drs. Linn and Huynh is provided in Appendix V. The discussions were carried out over an extended period of time with repeated exchanges of information, (for example, the latest response came May 15, 1983), underscoring the multiplicity of factors which need to be considered in any identification model and their correlated technical issues.

Based on the review of the literature completed in Activity 1 and the consultants' responses, four alternative discrepancy methods were analyzed in the paper, "A Comprehensive Analysis of Four Discrepancy Methods" (see Appendix W). The five sections of the paper contain: (a) a discussion of the general issues involved in defining learning disabilities with a statistical discrepancy model; (b) a presentation of four discrepancy methods for use in California; (c) an analysis of the alternative methods in terms of specified criteria related to measurement and application issues; (d) a comparison of the efficacy of the four methods when critical input factors (e.g., reliability coefficients, correlation coefficients, aptitude levels) are changed; and (e) the measurement experts' responses to specific questions. Since the intended outcome of Goal 2 was, in part, the specification of a discrepancy method for identifying LD students, this document was prepared to provide the critical information on which a decision could be made. Task Force II was charged with the task of recommending a particular method.

Activity 3: Selection of the Discrepancy Model Judged most Appropriate for The Community College Setting

Procedures. As stated previously, the Steering Committee recognized the need for a broad base of input from practitioners and experts. For Goal 2, a task force of practitioners was solicited by Laurel Best to review the discrepancy methods and the issues related to their implementation.

Task Force II was composed of 17 LD specialists and psychologists from community colleges. A list of task force members is provided in Appendix X.

As mentioned Task Force II members were charged to review the alternative discrepancy methods, evaluate these models' ability to deal with the essential factors, and to recommend a method for adoption. In preparation for this task, the members were provided with the documents prepared by the Steering Committee (Appendix W) and two published articles representing state-of-the art thinking. The intent was to facilitate an informed discussion when the Task Force met.

The Task Force and Steering Committee members met with Dr. John Poggio and Mr. Sid Cooley* on January 11, 1983, at Orange Coast College in Costa Mesa, CA. The topics for discussion were the alternative discrepancy methods, the trade-offs in using a discrepancy method, and the recommendations of the task force. The results of that meeting are provided in a report (Appendix Y). The conclusions to be drawn might be summarized as follows: (a) the alternative discrepancy formulas should be evaluated in a pilot and validation study; (b) a discrepancy score is not sufficient for determining that a student is learning disabled; (c) other selection criteria must be validated and included in an LD definition in addition to discrepancy and (d) continued research must be directed at developing assessment instruments and procedures without population biases, e.g., ESL, NES, LEP groups.

Extensive discussion focused on b, c and d (above) several of which were presented in the problems mentioned in the Goal 2 section of this paper. In addition, these issues will be discussed further as the selection model is described. In a sense, the results of Task Force II provide a summary statement of the intended outcome of Goal 2 as it relates to the establishment of a statistical discrepancy method for calculating and evaluating an aptitude-achievement discrepancy score. To accomplish Goal 2 the other criteria necessary in a selection model for LD students must be discussed.

---

*Mr. Cooley, a graduate research trainee at the KU-IRLD, played a major role in the preparation of the discrepancy materials.

Selection Model Methods

Definition of terms. At some point the reader is likely to be confused about the term "selection model" as used in this report. The term is used here in two ways. First, its general referent is to the identification and validation of those criteria other than discrepancy score information which would be used in determining a student's eligibility for services. This use reflects the recommendation of Task Force II and numerous consultants that (a) a discrepancy score cannot be validly determined for members of some populations, (b) a significant discrepancy score by itself is not indicative of a learning disability; and (c) other criteria should be examined when making that determination.

In its alternative form, the term "selection model" refers to the whole process of determining LD students as a selection process. Anastasi (1982) noted that in selection, each individual is either accepted or rejected. In the case of LD programs, the student is accepted or rejected for the LD program. "When selection is done sequentially, the earlier stages are often called 'screening,' the term 'selection' being reserved for the more intensive final stages" (p. 179).

In the broader context of community college recruitment and retention of students, the LD program represents just one possible student classification. Other classification examples include choice in the college curriculum as well as field of concentrated studies. Hopefully, this discussion provides a context within which the "selection process" of students into the LD programs can be more easily understood and emphasizes that the "selection process" involves numerous criteria, one of which is the aptitude-achievement discrepancy score.

Restatement of problems requiring selection model. First, as has been frequently stated, LD is a multidimensional problem and, although a significant aptitude-achievement discrepancy is one of its characteristic, in itself this is not conclusive evidence of a learning disability. Second, for some subpopulations within the community colleges formal assessment data are not available, usually due to lack of appropriate normative data on standardized test instruments. The Limited-English Speaking (LES) subpopulation is one such group. Third, in addition to specific characteristics of LD other characteristics exclude a student from being identified as LD. These criteria, which can be found in the "exclusion clause" of LD definitions, suffer from generally being too vague to be consistently meaningful. Fourth, all the criteria on which an eligibility decision is based have not been weighted and organized into a set of procedures which can be applied by the LD specialists and others charged with the task of determining student elibibility for LD programs.

The above four problem areas have been reviewed as a reminder to the reader as the methods of solving them are described in the methodology.

Concerns with standardizing the selection process. Several problems related to the LD eligibility criteria have been previously described. In one sense, they can be treated as the methodological problems and issues confronted by the field of learning disabilities, not exclusively the California community college system. However, some concerns expressed by practitioners in the community colleges and researchers need to be addressed because they impact on the utility of the outcomes being developed. These concerns have been expressed at various conferences, presentations of the project, and through correspondence. The concerns are not necessarily representative of any one group. However, representativeness has not always been the critical dimension on which decisions were made.

First, practitioners are concerned that changes in the eligibility criteria might lead to local decisions and more decisions based on an application of state level criteria. Some of the benefits to accrue from this include the need for uniformity across campuses, provision of equitable treatment of students, and the appropriate role of the professional in determining eligibility. While most would acknowledge there are inconsistencies in the current practices, it is questionable whether state-level criteria would undermine the local autonomy.

A second related concern is that student eligibility is currently based on the clinical judgment of the practitioner. Any comprehensive eligibility/criteria, however, could change this. For example, empirically derived criteria specify that a student is rated on each of the criteria and if he/she meets the criterion cutoff score the student is determined eligible. This selection strategy minimizes the LD practitioner's clinical judgment and effectively modifies his/her role.

A third concern stems from the multidimensional conceptualization of learning disabilities. Whether one is considering etiology, diagnosis, intervention or long-term consequences of learning disabilities, a multitude of perspectives present themselves. Consequently, it is debatable whether a proposed model can (a) be developed as a unitary identification model and (b) be fit to a statistically based identification model. Some might argue that given the low incidence figures associated with particular learning disability constructs, e.g., dyscalculia, dyslexia, dysnomia, etc., a statistical model is not realistic because the sample sizes needed for developing and testing such a model are not available.

The last concern to be addressed is a methodological issue concerning data collection. Numerous factors can bias the results of any field study and, thus, present sources of error in validity. For example, some researchers have suggested the possibility of a conflict of interest in instances when LD specialists whose position is dependent upon maintaining enrollment determine student eligibility. The same line of reasoning might also be applied to having LD specialists gather data for the project which, in turn, would be data used in establishing eligibility criteria.

Based on this brief overview, the reader is left to develop the pros and cons of the above issues. While the authors have avoided elaboration or comment at this time, these concerns did impact on the decisions made regarding the development of the selection process and the pilot study of the learning disability eligibility criteria.

Activity 1: Identify Factors That Are Potential Additions and Alternatives to the Discrepancy Model

Procedures. At the September, 1982, Chancellor's Regional Facilitators Meeting in Sacramento, the Steering Committee presented the goals, objectives, and activities of the Consortium project. Likewise, at the LD Division meeting of the CAPED convention in Los Angeles, Laurel Best, Don Deshler, and Daryl Mellard of the Steering Committee, made a similar presentation. The information obtained from this broad base of practitioners and administrators reinforced the complexity of the LD eligibility issues.

Specifically, these groups articulated concerns that the Consortium project consider how eligibility criteria might be applied to subpopulations (e.g., LES-NES, racial groups). Further, participants expressed concern that the "exclusion clause" of the LD definition is not defined well enough to be operational. Since neither of these issues, which were repeatedly expressed at subsequent meetings, e.g., Task Force II meeting, State Advisory Committee meetings, had received extensive discussion when the project was initially developed, it was necessary to expand the activities of the project in developing the selection model. In part, the results of this input solidified the notion of completing a pilot study in the first year of the project and later conducting a validation study of the model.

An additional source of information concerning the specification of factors for the selection model came from Dr. Lee J. Cronbach at Stanford Unviersity, an established leader in educational and psychological measurement. Dr. Cronbach met with members of the Steering Committee to discuss the problems relevant to determining eligibility criteria through a research design. The most significant conclusions from that discussion were: (a) clinical judgment not just a statistical model, is necessary for determining eligibility, (b) a model for providing a long-term solution to the problem of determining learning disabilities eligibility is lacking; and (c) the development of a selection model might begin by focusing on the outcomes of the treatment model. If the intervention goal for LD students is successful academic performance, one might begin to understand learning disabilities by studying the traits of students who respond and do not respond to the treatment (i.e., intervention). Based on such a study one might be able to evaluate a student for those characteristics and, in part, determine eligibility.

In summary, the results of this activity led to the specification of particular components for the selection model. Those components were (a) addressing the need for valid assessment of minority populations, (b) operationalizing the exclusion clause, and (c) developing a classification system for the community colleges that, in part, based eligibility in LD programs on how successful a student was as a result of the intervention. The resulting criteria might be applied with empirically derived cutoff levels and the clinical judgment of the professionals who are determining eligibility. The next task was to develop a plan that permitted an investigation of these components.

Activity 2: Selection of a Research Methodology

Procedures. In formulating the Selection Model, the importance of including factors relevant to such areas as student success or failure in the community college, interactions among student characteristics and classroom variables, as well as self-help and academic behaviors can not be denied. However, the task remained of determining which factors to include in the selection model. To this end two approaches were investigated: Naturalistic Inquiry and a synthesis methodology.

Naturalistic Inquiry (Guba, 1978) was suggested as a possible solution to this dilemma. Although a full discussion of the naturalistic inquiry research paradigm is beyond the scope of this report (see Guba & Lincoln, (1982) for a complete discussion of the Naturalistic Inquiry methodology), a brief description of its premises and methods will help the reader appreciate the initial appeal of this approach.

Essentially, the Naturalistic Inquiry design has the advantages of allowing researchers to study a problem the context in which it occurs without the constraints, assumptions, and influences imposed by experimental manipulation. Observers go into the setting, (in this case, regular and LD community college classrooms) and, using the techniques of observation, interview, and record review, collect notes on any factor that is even tangentially related to the topic under study. These notes are then studied, categorized, interpreted, and written up as detailed descriptions of the problem as it occurs within its natural setting. The premise is that through studying these detailed descriptions, the realities of the problem will emerge. The results will be based not on a prior assumption of reality but on the picture of reality which has emerged in the course of studying the problem in its own context.

Since so little is understood about learning disabilities as they occur and operate within the classroom, particularly at the community college level, Naturalistic Inquiry offered an attractive means for determining what factors relevant to students' functioning within the community college are significant in the referral and ultimate LD placement and success of students. Consequently, serious consideration of the naturalistic paradigm began. Dr. Thomas Skrtic, a researcher at the University of Kansas who was currently involved in a Naturalistic Inquiry study of special education in rural America, was consulted to help evaluate the feasibility of using Naturalistic Inquiry in the selection pilot study. As the data collection procedures were detailed under Dr. Skrtic's guidance, major disadvantages of using the naturalistic inquiry methodology surfaced. Foremost was the considerable cost involved in terms of time and training. For example, training sessions for all field workers would have to be conducted. This requirement alone presented numerous logistical problems, further complicated by the physical distance between Kansas, where the trainers were located and California, the location of the field workers and study sites. After the training, field workers would need to spend a considerable amount of time collecting data in order to get a full picture of each site. Once data were collected, the amount of time needed to organize, categorize, interpret and write up the case studies plus conduct member checks would take more time than that remaining for the completion of the study. Even paring down the number of sites in which data would be collected contributed little to making the naturalistic paradigm more feasible, given the constraints under which the project was operating. Although Naturalistic Inquiry was the preferred method for answering the questions under study, this approach was rejected in view of the limited time and resources was not possible.

Having abandoned the Naturalistic Inquiry paradigm, a synthesis methodology was formulated. Consequently, survey instruments were designed which made it possible to collect data from numerous sources thus setting up the opportunities for cross-checks similar to the member checks included in the naturalistic methodology. Although the surveys required that factors be pre-identified, a significant disadvantage of the survey approach, results of research studies conducted by the Kansas IRLD on the adult-aged LD population as well as the input of community college LD specialists on factors relevant to the success of LD students in California's community college setting offered guidance in selecting factors that would be powerful discriminators of successful LD (SLD), unsuccessful LD (USLD), and low achievers (LA). Also, by designing the surveys to that respondents were encouraged to add information they considered significant to the questions asked, factors otherwise overlooked by the preformulated survey questions could be identified. A complete description of the generation of selection survey instruments is provided in a later section of this report.

Activity 3: Selection of a Research Design

Procedures. As described, the clinical model for identifying LD suggested by Dr. Cronbach was based on serving those students who were judged most likely to be successful in the community college program and not serving students failing to show promise of success despite placement in the LD program. The appeal of such an approach to the identification of LD is that the effort and expense incurred in providing LD services has a payoff--the students who are served get better. By analogy, the students who demonstrated no apparent benefit from receiving the LD intervention would be no worse for having not been served.

A basic assumption of this clinical model is that the service provided within the LD programs is appropriate for eligible students. In reality, however, sufficient evidence does not exist to support such a claim. A standard treatment or cure for LD is not currently available. Each case is unique. Although the LD specialist does her/his best to plan and implement appropriate interventions, none guarantees in advance success for every student. In most cases, interventions evolve over the course of working with a student and success rates are, at best, variable.

With this in mind, the clinical model, i.e., identification of LD based on a high likelihood of success in the LD program, posed a nagging concern for those involved in designing the selection pilot study. The focal point of this concern was that just because a student was not successful in a LD program, it did not mean that he/she was not LD--only that the appropriate intervention had not yet been discovered. Indeed, it seemed likely that low achievers would be most likely to succeed in the LD program, given that their ability to learn is not handicapped. By setting such a criterion for identifying LD, it appeared probable that the most handicapped students would be the ones excluded from the very programs designed to help them by virtue of their own handicap and its resistance to remediation (DeLoach, Earl, Brown, Poplin, & Warner, 1981).

Still, it was important to know the factors which play a role in a student's success in the LD program, if only to increase a[illegible]ess of personal and programmatic factors which are beneficial to LD students. The solution that emerged from numerous discussions among members of the Steering Committee and project staff was to collect data on three subgroups of the community college population: (a) successful LD students, (b) unsuccessful LD students, and (c) low-achieving students. In general, low achievers were those who had been referred and then determined to be ineligible for LD service. To ensure that the LD students included in the samples were truly LD, a vital criterion was added to the subject selection procedures. That is, in the best professional judgment of the LD specialist, each selected LD subject must, without a doubt, be judged as LD. This criterion was added because of the high probability that some of the students served as LD might more appropriately be classified as something other than LD (DeLoach et al., 1981). Further discussion of the subject selection criteria and procedures is reserved for Goal 4 of this report.

Summary. The Selection Model discussed here evolved from the need for comprehensive analysis of student behaviors indicative of a learning disability. Since learning disabilities are not restricted to one population of community college students, and are not conclusively indicated by any one characteristic, multiple criteria must be developed on which subpopulation biases are known. Therefore, a research methodology was developed which permits the determination of empirical indicators of learning disabilities. In this way, criteria other than aptitude-achievement discrepancy score could be both statistically and clinically evaluated in determining a student's eligibility.

Summary of Goal 2

The purpose of Goal 2 was to specify a statistical model for evaluating aptitude-achievement discrepancy scores for statistical and educational significance and a selection model which evaluated other criteria for determining eligibility. As a result of Task Force II's efforts, a statistical discrepancy method was established for application in the pilot study. The Task Force recommended that a True Score Regression procedure be used in calculating and evaluating discrepancy scores. A research methodology was designed to investigate other criteria that might be useful for determining eligibility. This methodology and criteria will be discussed in Goal 4.

GOAL 3

Goal Statement: To establish the magnitude of discrepancy and/or cutoff levels necessary to verify a learning disability among community college students.

Problems addressed by Goal 3. Eligibility scores for student admission to LD programs must be established. Prior to PL 94-142 and the institutionalization of specific learning disabilities as a handicapping condition, the LD population consisted of a small, clinically defined group. However, as programs for the learning disabled have become more popular in recent years, the number of students referred for LD services has grown even more rapidly. Because of limited resources, decisions are now being made to establish cutoff levels for determining eligibility for such services. In the past, clinical judgment constituted a major identification strategy for LD populations. However, many of these clinical decision-making practices failed to systematically consider the values of key stakeholder groups (e.g., college administrators, LD specialists, Department of Finance personnel, etc.) in deciding the minimum eligibility criteria. The issue of setting eligibility cutoff scores is as central to the identification issue as the specification of which test instruments to use, for example. Appropriate cutoff levels can be established by using decision-making methodologies that explicitly consider the values, concerns, motives, and interests of the key stakeholders involved. In addition, as these procedures are based on both clinical and statistical criteria they can take into consideration the rates of correct and incorrect classifications, and assess the positive and negative consequences of any decision regarding cutoff levels.

The problem of establishing cutoff scores for eligibility to LD programs is multifaceted and, as previously stated, it cannot solely be based on statistics.

Along with statistical information about cutoff scores, decision-makers must consider social, educational, political, professional, and financial factors. To omit any one of these factors would be to ignore variables (e.g., financial) which could potentially destroy the viability of an LD program. Thus, the problem confronted in Goal 3 related to establishing a cutoff score using procedures which recognize the diversity of factors relevant to the issue and the potentially competing values of the persons impacted by the decision.

Statistical procedures and decision theory can be applied to solve such problems as determining eligibility scores or cutoff levels when many diverse factors and competing viewpoints must be considered. The advantage of such procedures is that all the alternative solutions can be considered and the criteria ultimately selected will reflect the values of those involved.

In this problem, the various viewpoints were directed at evaluating the relative merits of four possible outcomes of determining a student's eligibility for an LD program. These four outcomes were: (a) identifying as LD a true LD, (b) identifying as LD a true non-LD, (c) not identifying as LD a true LD, and (d) correctly rejecting as LD a true non-LD.

In essence, the procedures involved having those stakeholders concerned with the cutoff level decision identify their values about the decision and numerically represent the importance of the four outcomes according to their values. Thus, for each group, the four outcomes represent different utilities. For example, community college classroom instructors may have different values than Department of Finance representatives about identifying LD students. Such differences can be translated into an understanding of how each group views the consequences of the four possible outcomes. This information was obtained on a Multiattribute Utility Questionnaire. These utilities are used in determining a particular degree of certainty, i.e., probability level, that must be satisfied to further consider a student learning disabled. For example, aptitude-achievement discrepancy score was described in Goal 2 as one component of determining a learning disability. Based on the results of Goal 3, one can determine whether or not a student's discrepancy score is severe enough to meet the degree of certainty required to be further considered. Assuming that the stakeholders indicated that their values are served by a 90% level of confidence (9 out of 10 times), then the student is correctly classified. If one knows how frequently a particular discrepancy score occurs in the normal population and the LD population, one can determine whether or not the score will result in 90% correct classifications.

Intended outcomes. Goal 3 and its related activities were formulated to provide a single, specific outcome: a level of student performance from which the verification of a learning disability can be made on a consistent and agreed-upon basis.

Major activities. Three major activities were completed as part of this goal in the framework of decision theory: (a) formation of Task Force III; (b) preparation of example lists of values; and (c) implementation of utility theory with Task Force III members.

Method

Participating consultants. As indicated on the Organization Chart (see Introduction Section), Drs. Detlof von Winterfeldt, the University of Southern California's Social Science Research Institute, and Hutton Barron, the University of Kansas School of Business, provided valuable consultative assistance. Dr. Barron worked with the Steering Committee members at the University of Kansas and provided extensive background information about the decision-theoretic model utilized in the procedures.

Dr. von Winterfeldt was requested to participate as a consultant because of his recognized international leadership in the application of decision-theory techniques to establishing public policy. His work has included such diverse topical areas as land use management, energy development, and building structure codes. Dr. von Winterfeldt had primary responsibilities for organizing, completing, and analyzing the activities for Task Force III. Appendix Z provides a formal technical statement of the problem as formulated by Dr. von Winterfeldt.

Activity 1: Formation of Task Force III

Procedures. Steering Committee members and members of the Chancellor's Office met to identify generic groups for determining eligibility standards for the community college LD program. The following is an exhaustive list of those groups: (a) community college LD students; (b) community college LD specialists; (c) community college administrators; (d) parents of LD students and other advocacy groups; (e) community college counselors; (f) learning and resource center teachers; (g) content teachers (non-special educators); (h) state agencies; (i) researchers in learning disabilities; (j) nonhandicapped students; (k) national LD experts; (l) ancillary service personnel; and (m) legislative personnel.

From the above list of key stakeholders, specific agencies and organizations were named along with individuals who might have values similar to other members within that agency. These individuals were subsequently contacted by the Chancellor's Office, and requested to participate. Appendix AA contains a listing of these specific agencies and individuals asked to participate in the project. These participants composed Task Force III.

As noted in the Appendix, the above generic groups were combined to form seven larger groups. The logic behind such a grouping was partly organizational, but primarily it is intended to provide the greatest homogeneity of the values, concerns, interests, and viewpoints of those involved. For example, school psychologists, speech clinicians, and counselors were considered as likely to have similar values concerning the identification of LD students. Based on this similarity of concerns, they were grouped together as ancilary services. Similarity of basic values, however, did not imply that members weighed the consequences of identifying or misidentifying students identically or that the values of the groups were ranked in the same order. For example, two values of these ancillary services groups might be: (a) meeting legal requirements and (b) role effectiveness. For the school psychologists, meeting legal requirements may be a more important value than role effectiveness. However, for the speech clinicians, these values may be reversed in order of importance. Thus, similarity in values does not imply that they were considered of equal importance by stakeholder group members.

Three of the generic groups, national LD experts, LD researchers, and parents, were eliminated from further consideration due to (a) their values were represented by others, (b) there was difficulty in identifying individuals as participants, or (c) they were not considered significant for inclusion.

Activity 2: Preparation of Example Lists of Values

Procedures. Since this was the first instance of applying a decision-theoretic approach such as utility measurement to a special education issue, i.e., establishing an eligibility score, it was learned that the participants might be too unfamiliar and uncomfortable with the task requirements to participate effectively. To facilitate the process it was decided to establish an example set of values for each respective stakeholder group. The initial task of developing these value lists was completed at the LD Division meeting of the CAPED convention (October 21 in Los Angeles).

Those attending the meeting were divided into small groups and asked to assume the role of various stakeholder groups (e.g., counselors, LD students, college administrators). Each group was asked to identify the values, concerns, motives and viewpoints that would be representative of the respective group (Appendix BB). For example, those in the role of college administrators listed the following concerns relevant to the "misses" in LD classification decisions, i.e., classifying a true learning disabled student as non-learning disabled. From an administrator's point of view, therefore, a "miss" classification could result in a higher dropout rate from college, faculty complaints, failure to meet legal compliance requirements, etc. In a similar manner, a tentative list of values was developed for each of the seven stakeholder groups.

Next, the lists were organized into unique statements of values for the seven stakeholder groups and evaluated for how well the particular values are impacted by correct and incorrect classifications of students as LD or not LD (Appendix CC). For example, one potential concern (i.e., value) of state level administrators was overall cost of LD programs. This value has a positive or negative, or zero impact on correct and incorrect classification of students.

This evaluation of the procedure for determining the impact of classification decisions was reviewed by Steering Committee members and four LD specialists from the community colleges. The results of the review were used in developing the Multiattribute Utility Measurement Questionnaires and the materials used by Task Force III.

Activity 3: Implementation of Utility Measurement with Task Force III Members

Procedures. On March 24, 1983, Task Force III met under the direction of Dr. Detlof von Winterfeldt and his assistant, Mr. Greg Griffin, for the purpose of determining the eligibility criterion (i.e., the discrepancy score) for verifying a learning disability. Representatives (a total of 20) for each of the seven stakeholder groups and Steering Committee members also attended (Appendix DD). One participant judged that it was inappropriate to participate as a stakeholder group member and withdrew (Appendix EE).

Each stakeholder group attended meetings during which the utilities of correct and incorrect classifications were assessed. The participants were briefed on the overall goals, procedures, and current activities of the project. For many participants, this was the first formal participation in the project; for others the meeting provided an opportunity for review and updating of the activities.

The following sequence of events was followed with the stakeholder groups: (a) introduction and statement of the problem; (b) general description of the planned activities; (c) completion of the Multiattribute Utility Questionnaires which included reviewing and labeling stakeholder group value statements, evaluating the desirability and undesirability of classification decisions (e.g., hits, correct rejections, false alarms, and misses) for each value, and weighting the relative importance by each group of their values (Appendix FF); and (d) completion of a followup questionnaire (Appendix GG).

Specifically, the task related activities of setting an eligibility score as described in Appendix GG can be summarized as follows. First, each stakeholder group identified their concerns (i.e., values) about the outcomes of classification decisions regarding learning disabled and non-learning disabled students. The four outcomes were: (a) hit (classifying as learning disabled someone who truly is), (b) misses (classifying as non-learning disabled someone who truly is LD), (c) false alarms (calling someone learning disabled who truly is not) and (d) correct rejections (classifying someone as non-learning disabled who truly is non-LD). Second, for each concern, the stakeholders located the four outcomes on a numerical scale relative to each other. The best outcome was scored 100; the worst outcome received a zero rating. The other two outcomes were located on the scale so that the relative distances were meaningful. Last, the participants rank-ordered their values in terms of importance. The value thought to be of most overall importance for classifying students received the greatest weighting.

Results. The results of the analysis are presented in Appendix HH. Interpretation of the desirability of the four classification options suggests that "hits" and "correct rejections" are the most valued outcomes across all groups. This indicates the importance the groups place on making the correct decisions relative to any student. When such emphasis is placed on correct decisions, there is less concern for the consequences of either "misses" or "false alarms."

For the seven stakeholder groups combined, the likelihood ratio is 13.4 to 1 when the incidence level is 5%. In other words, when 5% of the population is thought of as learning disabled, the stakeholders want to have 93% confidence in their classification decision before verifying a learning disability on the basis of an aptitude-achievement test discrepancy score.

The three groups setting the strictest critera for eligibility were the Department of Finance (76:1), the Legislative group (23:1), and the College Administrators (17:1). The most lenient criterion were proposed by the Non-LD Instructors (5:1), Low Achiever Representatives (6:1) and LD Specialists (6:1). The numbers in parenthesis represent the likelihood criterion that must be met to identify a student as learning disabled.

Results of exit questionnaire. Seventeen of the 20 Task Force III participants completed a questionnaire (Appendix II) to indicate their reactions to the Goal 3 activities. The results of this survey are found in Appendix HH. In general, the respondents' statements about the session and the activities were positive. In particular, Task Force III members indicated that they understood the use of the value statements, the weighting procedures for ranking their values, and scaling the four outcomes from best to worst. In comparison with the rankings given the preceding areas, they felt somewhat uncomfortable making the judgments themselves.

Limitations. The examples of values (Appendix DD) developed in Activity 2 for this elicitation session apparently were adequate, since most of the stakeholder groups responded to the value statements with little alteration. However, other groups altered one or more value statements. In retrospect, the prior preparation of the value statements may not have been necessary.

The problems observed using this methodology were minor. In retrospect, two other additional changes seem desirable for any repetitions of the procedure. The first would be to provide a fuller description of the procedures to the participants prior to the meeting. Several individuals commented that such an advance organizer would have been helpful. The second change would be to provide an opportunity after the elicitation session for general debriefing and questions. Thus, some of the participants commented that they were still uncertain about how the results of the elicitation session would lead to a cutoff score or how the procedures would be integrated.

Summary. The present project represents the first reported instance of following a formalized decision-theoretic approach to establish a criterion score for LD programs eligibility. In short, the procedure quantified the potentially competing values of various stakeholder groups and the utility of classification decisions based on those values. Such an approach was thought necessary for numerous reasons including the need to provide consistent, clearly defined criteria for the numerous community college settings in California. The results of the goal-related activities indicate that Task Force III participants would recommend eligibility criteria setting the odds ratio of LD as approximately 14 to 1. Yet to be determined is the distribution of discrepancy scores for the LD and non-LD Community College populations for each of the aptitude and achievement test pairings used to calculate the discrepancy scores.

GOAL 4

Goal Statement: To empirically validate the certifying properties (attributes, models, and discrepancy or cutoff levels) of a learning disability as it exists among community college students.

Problems addressed by Goal 4: Empirically validated data on existing identification procedures are lacking. To date, no comprehensive studies have been conducted within the learning disability field to validate current identification practices. The consequences of this are manifold, foremost among them being the inability of most LD programs, even after years of providing LD service, to justify their placement and intervention decisions with much more than an appeal to under-achievement. Likewise, the efficacy of the model proposed here remains unknown until the LD identification procedures and eligibility criteria developed in Goals 1, 2, and 3 of this project have been rigorously tested and refined through empirical validation.

For these reasons, it was imperative that a thorough field-test of the proposed procedures be conducted in the California community college system. According to the original plan a validation study was to have been completed during the current (1982-83) academic year. However, as the complexities of the identification issues became evident, the Chancellor's Office recommended that a pilot study rather than a full-scale validation effort be conducted. This recommendation was made to ensure that efforts to accomplish Goals 1, 2, and 3 would be sufficiently comprehensive and rigorous.

Intended outcomes. The California community college system was in the unique position of being able to systematically gather the data necessary for validating its selection procedures and empirically determining the requirements and consequences of their implementation.

A pilot study conducted in selected community college learning disability programs provided the groundwork for a more extensive validation at a later date and allowed potential problems in the identification procedures to surface and be resolved prior to the full-scale validation study. The following intended outcomes were identified as relevant to the pilot study:

1. Establish a working relationship and credibility with the involved schools.
2. Initiate preliminary data analysis (e.g., descriptive statistics comparing scores of LD, low achievers, (LA), and non-handicapped (NH) students on each discrepancy/selection instrument).
3. Evaluate the utility benefits of the discrepancy/selection identification procedures.
4. Determine the feasibility of the discrepancy/selection identification procedures (by obtaining evaluation data from participating LD specialists).
5. Establish an evolving data base for a proposed normative and validation study.

Major activities. The pilot study involved the implementation of procedures and instruments described in Goals 2 and 3, which were developed as parts of the discrepancy and selection models for identifying learning disabled community college students. The study consisted of two major activities: The discrepancy model pilot study and the selection model pilot study.

Discrepancy model pilot study. As described previously, the discrepancy model for LD identification is based on an educationally and statistically significant psychometric discrepancy between a student's intellectual ability and academic or achievement in one or more educational areas. Therefore, the purpose of the discrepancy pilot study was to apply (a) the True-Score Regression discrepancy model recommended by Task Force II, (b) the recommended IQ and academic achievement assessment instruments, and (c) discrepancy cutoff points to existing populations of LD, LA, and NH community college students to determine how the discrepancy model would differentiate these subgroups of community college students.

In addition to evaluating the utility of the discrepancy model, the study was designed to determine the feasibility of using the process to identify LD in California's community colleges. To aid in the selection of the best discrepancy formula for use in the community colleges, several discrepancy formulas (Appendix JJ) other than the True-Score Regression formula were applied to the data and subsequent results were compared.

Selection model pilot study. The selection model for the identification of learning disabilities is based on the premise that an ability/achievement discrepancy is not sufficient to diagnose LD, that is, a true learning disability is best identified in terms of a cluster of behaviors which includes a discrepancy score in addition to other student behaviors that do not lend themselves to standardized ability/achievement testing. These behaviors include the student's response to instruction and to the learning environment, self-help behaviors, sociability, self-attributions, history of schooling, etc.

The purpose of the selection pilot study was to field-test a set of survey instruments designed to tap differences among the various subgroups of community college students. As in the discrepancy pilot study, the utility as well as the feasibility of the selection procedures were the primary concerns.

The following sections provide the procedural details and the results of applying the two models.

DISCREPANCY PILOT STUDY

Method

Activity 1: Selection and training of Discrepancy Model Pilot Study Task Force.

Procedures. A Pilot Study Task Force comprised of LD specialists in 25 California community colleges was created to pilot the identification procedures developed in Goals 1, 2, and 3 as well as to provide formative evaluations of the statistical discrepancy model. As in the formulation of the task forces for the other project goals, volunteers for the pilot study task force were solicited by Laurel Best. In locating volunteers, Best attempted to include LD specialists from all areas of California in order to obtain a sample which would be representative of students in all types of community colleges--urban, suburban, and rural; large, medium, and small student populations; single-campus or multi-campus sites. Individuals who agreed to participate in the study were given the choice of collecting data for the discrepancy pilot, the selection pilot, or both. Fourteen of the twenty-five participating LD specialists elected to collect data for the discrepancy pilot. Within this group, four volunteered to collect data for the selection pilot as well. A list of the participating community colleges and LD specialists responsible for the data collection is included in Appendix KK.

Since collection of data for the discrepancy pilot did not involve unfamiliar or complicated procedures, specialized data collection training was not necessary. Instead, a set of detailed written instructions for collecting the necessary data was prepared and presented to the participating LD specialists at a meeting on February 9th and 10th, 1983. The discrepancy pilot study materials (described in a later section of this report) were described and explained in detail and questions were answered as they arose. Further training was not necessary. Once the data collection began, LD specialists were encouraaged to contact Daryl Mellard or Ron Baker in case problems would arise. However, the only major problem brought up related to difficulty in meeting the April 15th data collection deadline due to problems with scheduling all the testing sessions.

Activity 2: Selection of discrepancy assessment instruments.

Procedures. The discrepancy pilot study consisted of administering an intellectual aptitude test and an academic achievement battery to the groups of community college students targeted for the study. Because of the diversity of tests currently used in California's community college LD programs (Ostertag & Baker, 1982), it was critical that a set of valid and reliable assessment instruments be specified and that only those instruments be used to collect the needed aptitude/achievement data. The reader is reminded that the technical adequacy of the most commonly used LD identification instruments was evaluated by a group of assessment experts as part of Goal 1. Based on their recommendations, a set of acceptable assessment instruments was identified for use in the discrepancy pilot. For each subject selected for the study, one aptitude and one achievement measure was to be administered. Choices for the aptitude battery included: (a) Woodcock-Johnson Psychoeducational Battery, Part I: Cognitive Ability Tests; (b) Wechsler Adult Intelligence Scale: Revised Edition; or (c) Wechsler Adult Intelligence Scale. For the achievement battery the choices were: (a) Woodcock-Johnson Psycho-educational Battery, Part II: Achievement Tests; and (b) the Wide Range Achievement Test.

Activity 3: Preparation of materials for data collection.

Procedures. One major problem in attempting to collect data in California when the designers of the research procedures and instruments are located in Kansas was that once the LD specialists began collecting data, no one would be available on site to answer questions about the data collection. For the discrepancy pilot study, subject selection and management of data recording represented the two most difficult procedural aspects. Recognizing that the success of the pilot study relied on these two components, Institute staff developed a set of specific data collecting materials which anticipated the types of problems which could arise and clearly described how to handle them.

Before the data collection procedures were finalized, a description of the activities considered necessary for collecting the aptitude/achievemeet data on the three subgroups of community college students was drafted and sent to members of the Pilot Study Task Force (See Appendix LL). Members were asked to review the procedures and to suggest ways to improve them and to identify those procedural aspects which would be difficult to carry out.

A number of Task Force members responded with suggestions and concerns most of which focused on locating subjects, especially low-achievers, and the limited time available to administer the assessment instruments (These responses are included in Appendix MM). In formulating the final set of procedures for the discrepancy pilot, the Task Force members' suggestions were considered and implemented whenever possible.

The final set of data collection procedures and materials was developed on the basis of the feedback provided by the Task Force members. Thus, directions for implementing the pilot study were thoroughly explained in a 12-step procedural description detailing the LD specialist's responsibilities in carrying out the study from the initial inspection of the pilot study materials to mailing the completed research materials at the end of the study.

In addition to the detailed procedural description, a checklist of tasks involved in the discrepancy pilot was developed to facilitate management of the data collection. The LD specialist was to record the starting and completion dates for each task as well as to estimate the total time necessary to complete each task. This information would later be used to judge the feasibility of the proposed procedures for LD identification. Space for recording notes and comments was included on the checklist. It was not mandatory that each LD specialist proceed through the data collection activities in the order specified by the otuline of tasks. Consequently, LD specialists were instructed to make adaptations as they were necessary to accommodate their individual situations. The procedural description and the outline of tasks are included in Appendix NN.

In addition to the Procedural Description and Outline of Tasks, a packet of color-coded research materials was prepared for each research site. Data recording forms were printed on a separate color for each individual group from which data were to be collected--blue for LD students, pink for LA students, golden rod for NH students, and green for LD Specialists. This color coding was designed to assist in data collection. Figure 1 contains a listing of the packet contents.

FIGURE 1

Contents of discrepancy model data collection packets

1. Student Consent Forms (Form B)--10 copies, 3 blue, 3 pink, 3 golden rod, 1 green;
2. Stamped envelopes for mailing consent forms to students--9 envelopes;
3. Mailing envelope for returning completed research materials;
4. Student Data Sheets (Form C)--14 copies, 3 blue, 3 pink, 3 golden rod, 5 extra for recording data on non-selected students;
5. List of random numbers
6. Formative evaluation sheet (Form D)--1 green;
7. Record of Student Refusals to participate;
8. LD Specialists Checklist of Tasks (Form M)--1 green.

The packets of research materials were distributed to the participating LD specialists at a Task Force meeting held in Cypress College or DeAnza College on February 9th and 10th, 1983. Participants were given their choice of attendin either location for training.

Activity 4: Selection of research subjects.

Procedures. Three groups of students were targetted for participating in the discrepancy pilot study: 1) LD students; 2) LA students; 3) NH (non-handicapped) students. The following criteria were employed for identifying eligible students from each of these groups.

A. LD students

1. Must meet the following definition:

   A specific learning disability refers to disorders in which an individual exhibits a significant/severe discrepancy between the current level of developed intellectual abilities and academic performance despite regular instruction and educational opportunity, as currently measured by professionally recognized diagnostic procedures. Academic performance refers to achievement in the following areas: listening comprehension, oral expression, written expression, basic reading skills, reading comprehension, mathematical calculation and mathematical reasoning. Specific learning disabilities are often due to constitutional, genetic and/or neurological factors and are not primarily due to: visual (blind) or auditory (deaf) sensory deficits, motor handicaps, severe emotional disturbance, environmental or economic disadvantage, cultural/language difference or mental retardation.

2. The LD specialist should have no doubt that this student is learning disabled;

3. The students should have spent a year or less in the CC LD program. If you select students who have been in the LD program for more than one year, select those who have not responded to your intervention.

4. The students must be enrolled in or have completed at least 8 credit hours.

B. Low-Achieving Students

1. The students were referred to the LD program as having possible learning disabilities;

2. As a result of your standard evaluation procedures (preferably a comprehensive evaluation), it was determined that the students did not meet the criteria of the LD program;
3. The student's low achievement was not attributed to developmental disabilities, mental retardation, emotional disturbance, a physical handicap, or sensory deficits;
4. The students are best considered as "low-achievers";
5. The students must be currently enrolled in or have completed at least 8 credit hours.

C. Nonhandicapped students

1. Based on self-report, these students must never have participated in a special education program or have been referred to one;
2. The students must currently be enrolled in or have completed at least 8 hours;
3. All forms relevant to this group are color-coded with the color GOLDEN ROD.

The eligibility criteria for LD group included the requirement that in the LD Specialist's best clinical opinion the individuals selected were truly LD. This restriction was included because of the problem inherent in assuming that past classification as LD is sufficient evidence that a student actually fits that identification category. Given the recognized problems of identifying LD within the community college population, the inclusion of this requirement represented an important safeguard since it is likely that not all students currently served as LD are truly LD. Requiring that each subject meet a set of standard, observable criteria as well as the more intuitive clinical judgment criteria typically used by LD teachers when diagnosing LD - increased the likelihood that the subjects within the criterion groups were five representatives of the subgroup to which they had been designated. Because the evaluation of the discrepancy model's utility for differentiating LD, LA, and NH students relies on the accuracy with which initial classifications were made, this additional criterion represented an important check against developing procedures which would accurately identify those who are currently placed in LD programs as opposed to those who are truly learning disabled.

The final selection of participants from each of the three groups was based on a random selection procedure using a list of random numbers provided the pilot study materials packet. For this purpose the LD specialists were instructed to locate comprehensive rosters of each of the three groups of potential subjects:

1. students enrolled at the community college;
2. LD students currently enrolled in the LD program;
3. LA students who had been referred, evaluated but not placed in the LD program.

The specific directions for identifying subjects for each group are contained in the discrepancy study procedural directions included in Appendix NN. A summary of the major components of the selection process are outlined below.

Once they had located the three rosters of potential subjects, the LD specialists were directed to use a random selection procedure to select the final group of subjects. Essentially, this process required the LD specialists to select as subjects those students whose ordinal position on the roster matched the random number taken from the random number list. This matching process was used until three students who met all the selection criteria for that group were identified. LD specialists were advised to select several alternate subjects in case those initially chosen did not agree to or were not able to participate in the study.

After selecting subjects and alternates, the LD specialists were to personally contact each potential subject to solicit his/her participation. An outline of the information to be presented in this initial conversation was provided (see Step 5 of the discrepancy pilot study procedural description in Appendix NN). Briefly, the LD specialists' introduction of the project to the subjects included: (a) a description of the project as a study of the influences that lead to a student's success and lack of success in the community college; (b) an explanation of the student's role in the project as allowing the teacher to look at some of the skills, such as reading and math, which the student had acquired during his/her education; (c) an estimation of time requirements as approximately 2 hours; (d) an assurance of complete confidentiality (scores would be reported as group scores rather than individual scores and course grades would not be affected by a student's decision to participate or not participate in the study). A record of refusals and reasons for refusals to participate was maintained on all contacted students (Appendix OO).

After a student had agreed to participate, he/she completed and signed a consent form (Appendix PP). This step was typically completed at the time of testing. A duplicate copy of the permission form was provided for the student.

Activity 5: Collection of aptitude/achievement data.

Procedures. Once the subjects from each of the participating community colleges had been selected, the LD specialists initiated data collection in two ways--test scores were taken from existing student records and students were individually tested using the required assessment instruments. In some instances, part or all of the necessary test score information was already available in existing records. LD specialists were directed to check each subjects existing records and record old test scores if these met the following criteria: (a) scores were taken from the aptitude and achievement tests specified for use in the discrepancy pilot study; (b) testing had been carried out after February 1, 1982; (c) testing was judged to be a valid and reliable measure of the student's performance; (d) testing was individually administered.

For subjects who did not have suitable test scores available or who needed to have certain parts of the assessment battery administered, LD specialists scheduled individual appointments to administer the necessary tests. Standard administrative procedures described in the test manuals were used. Testing could be conducted in two sessions, if necessary, however, the aptitude assessments were always to be completed first.

An individual Student Data Sheet (Appendix QQ) was completed on each student after the testing session and record review. The data sheet consisted of questions related to demographics, school history, and self-rankings on areas of academic difficulty as well as space for recording the aptitude and achievement test scores. An ID code for the student, school and group assignment (LD,LA, NH) was also recorded.

Although the discrepancy pilot study only required that test data be collected on three LD, three LA, and three NH students at each participating community college, LD specialists were encouraged to collect data on additional students if such were already available. Student Data Sheets were specially coded for these students.

Activity 6: Formative evaluation data collection.

Procedures. In addition to the discrepancy test score data collection, participating LD specialists were asked to evaluate the procedures/instructions provided for them in conducting the discrepancy pilot study. This information would be used in revising the discrepancy procedures prior to initiating the comprehensive validation study scheduled for the following year.

Each discrepancy pilot study material packet contained a Formative Evaluation Sheet --Form D (Appendix RR). For each step of the study, LD specialists were asked to provide evaluative comments on clarity of instructions, effectiveness of procedures, and major problems. The completion of the Formative Evaluation Sheet terminated the data collection steps of the discrepancy pilot study. At this time all materials were to be returned to Ron Baker.

Data analysis. Data were received from all 15 participating colleges. However, the amount of data received from each campus varied, therefore, specification of sample sizes and specific data will be described under results.

The data analysis consisted of tabulating descriptive statistics on marker variables, e.g., age, sex, race, etc., and measures of academic performance. Lastly, the aptitude-achievement test scores were used in calculating discrepancy scores for the three subject groups: the learning disabled (LD), the low-achiever (LA), and the randomly selected (RS) sample.

Results

Description of Samples

The LD population is represented in two samples. The first (N = 40) was randomly selected according to the procedures described earlier whereas the second group of LD students (N = 11) includes students for whom the LD specialists had already existing test data, e.g., aptitude and achievement scores. These data were collected to further describe the LD population of the community colleges. Since little practical difference was found between the two LD samples the two LD groups were combined for the analysis of discrepancy scores.

The low-achiever sample included 30 students, and randomly selected sample 36 students. If each LD specialist had provided data on three subjects per group, as requested, the sample sizes would have been 45 for each group and thus yielded a more representative and stable picture of the three groups.

Table 2 provides the frequency and percentage distributions of the students on sex and race variables. The frequency distribution for the community college was necessary to determine how closely those population characteristics are represented in these sample groupings. On the sex variables, both sexes appear to be equally represented. Although there are nearly the same number of males and females in three of the groups, the non-randomly selected LD group contains nearly three times as many males as females. On the race, ethnic group variable, the majority of the sample is Caucasian. Blacks and Hispanics are also represented across most of the groups.

Table 2

Sex and Race Frequency Distributions for Student Samples

Subject group		LD	LA	RS	N RLD
Sex	Males	21(52.5)	13(43.3)	14(38.9)	8(72.7)
	Females	19(47.5)	17(56.7)	22(61.1)	3(27.3)
Race/ Ethnic Group	Caucasian	29(72.5)	15(50.0)	28(77.8)	8(72.7)
	Black	5(12.5)	4(13.3)	3(8.3)	1(9.1)
	Hispanic	6(15.0)	8(26.7)	0(0)	1(9.1)
	Asian	0(0)	1(3.3)	1(2.8)	0(0)
	Other	0(0)	2(6.7)	4(11.1)	1(9.1)
Total sample size		40(100)	30(100)	36(100)	11(100)

Number of Schools Attended

Since school failure is often associated with poor instructional programs, the number of times a student changes school may be an indicator of educational consistency. Table 3 presents the frequency and corresponding percentage of students who attended one or more elementary, junior-high (including middle school), and senior-high schools. The percentages are enclosed in parenthesis. The data show a strong consistency across the four groups and the three levels of education. Thus, the pattern indicates that for all four groups at least half the individuals changed elementary schools. At the junior-high and senior-high school levels, at least 63% of the sample remained in the same setting. Until contradictory data are obtained, present data suggest that changes in school settings are not the basis for low achievement among LD and LA. This finding, however, does not rule out the possibility of some differential effect of changing schools for certain portions of the population. Such an interaction with LD and other low achievers should be considered as a possibility.

Grade Retentions

One indicator of previous academic performance is the frequency with which an individual was retained at a particular grade level. Table 4 provides the frequency with which students in the four samples were retained. Percentage values are enclosed in parenthesis.

Table 4

Frequency of Student Retentions

Frequency	LD	LA	RS	NRLD
0	20(50)	24(80)	34(97)	6(67)
1	18(45)	4(13)	1(3)	2(22)
2	2(5)	2(7)	0	1(11)

Table 3

Frequency Data on Changes in Attendance Centers

	LD			LA			RS			NR LD		
Frequency	Elem	Jr.High	High	Elem.	Jr.High	High	Elem.	Jr.High	High	Elem.	Jr.High	High
1	20(50)	30(77)	25(63)	16(53)	24(83)	19(66)	18(50)	23(64)	27(75)	4(40)	7(70)	8(73)
2	10(25)	8(21)	10(25)	8(27)	3(10)	5(17)	10(28)	10(28)	6(17)	2(20)	2(20	2(18)
3	7(18)	0	2(5)	2(7)	1(3)	1(3)	4(11)	0	2(6)	2(20)	1(10)	1(9)
4	1(3)	0	0	0	0	0	2(6)	1(3)	0	1(10)	0	0
5	2(5)	0	0	2(7)	0	0	0	0	0	0	0	0
6	0	0	0	2(7)	0	0	1(3)	0	0	1(10)	0	0

Although the data indicated that students in the four groups changed schools with approximately the same relative frequency, the number of grade level retentions indicates differential group effects. That is, among the LD students, 50% had been retained in school, compared to only 20% of the low achievers and 3% of the randomly selected students. This finding suggests that the LD students have a history of academically related problems. One indication of the severity of these problems was the frequency of the grade level retentions.

Primary Area of Academic Difficulty

Eligibility for LD intervention services is determined by achievement in seven academic areas: listening comprehension, reading recognition, reading comprehension, math calculation, math reasoning, written expression, and oral expression. That is, a student must demonstrate a significant discrepancy in one or more of these areas to be eligible for LD service. As part of the collected data, students were asked to identify one primary area of difficulty and any number of other secondary areas. Since only one primary difficulty area was indicated the sum of the percentages should equal 100% (See Table 5). However, students could indicate multiple responses to other areas of difficulty. Therefore, for any one area, for example, listening comprehension, the percentage indicates the percent of students who considered it an area of difficulty. Consequently, the percentages sum to a value greater than 100%.

Among the LD students sampled, the primary area of difficulty was reported to be written expression. Thus, 41% of the LD sample and 36% of the non-random LD sample ranked written expession as **the** major area of difficulty. Assuming that the curricular demands are the same for all community college students, it is interesting to contrast the responses

Table 5

Frequency of Primary and Secondary Areas of Academic Difficulty

Area	LD		LA		RS		NR LD	
	Primary	Secondary	Primary	Secondary	Primary	Secondary	Primary	Secondary
LC	5(13)	12(30)	4(14)	14(47)	2(7)	0	1(9)	3(28)
RR	11(28)	15(38)	0	5(17)	1(4)	2(6)	2(18)	6(55)
RC	4(10)	21(53)	9(31)	10(33)	4(14)	2(6)	2(18)	4(36)
MC	1(3)	22(55)	2(7)	10(33)	8(29)	4(11)	1(9)	2(18)
MR	2(5)	16(40)	3(10)	9(30)	6(21)	5(14)	0	4(36)
WE	16(41)	17(43)	7(24)	13(43)	6(21)	5(14)	4(36)	6(55)
OE	0	9(23)	4(14)	5(17)	1(4)	3(8)	1(9)	4(36)
Total Frequency	39	112	29	66	28	21	11	29
Ratio of Areas to Sample Size	-	2.8:1	-	2.2:1	-	.56:1	-	2.5:1

of the LA and random sample with the LD students. For the random sample, the areas of math calculation and math reasoning were ranked at least as frequently as written expression. Among low achievers, reading comprehension was cited most frequently, with written expression responded as second in the frequency ranking. These findings must be replicated before their diagnostic value can be concluded. However, support is indicated for previous research citing written expression skill as an effective discriminator among LD, LA, and normal students (Warner, Alley, Deshler, & Schumaker, 1980).

When the subjects were asked to identify other areas in which they experienced difficulty, a great deal of variation was found in the number of areas identified across the four groups. For example, the ratio of number of areas checked to the number of students in the sample groups was 2.9, 2.2, .56 and 2.5 to 1 for the LD, LA, random sample and non-random LD student groups, respectively. In other words, on the average, each student in the LD group checked almost three (2.8) areas. The non-random LD students checked a similar number of areas, 2.5. The LA had checked nearly as many areas, 2.2. In contrast, almost half of the random sample of students did not check any additional areas (ratio .56 to 1).

Secondary areas of difficulty varied greatly among the groups. The LD sample, for example, cited most frequently math calculation (marked by 55% of the sample); the least cited area for this group was oral expression (marked by 23%). For the non-random LD sample, the most frequent secondary problem was a tied ranking between reading recognition and written expression (55%); the least frequently cited area was math calculation (18%). Among the LA students, listening comprehension was the most consistent problem as indicated by 47% of the sample. The areas in which these subjects experienced the least difficulty was reading recognition and oral expression (17%). Although few of the random sample of students listed any secondary problem areas, the most frequently cited areas were math reasoning and written expression--both marked by 14% of the sample. No subjects indicated any problem in listening comprehension. Again, the above percentages and frequencies must be interpreted based on the recognition that the sample sizes are small and, therefore, perhaps less likely to be reliable indicators of the actual populations they characterize.

Aptitude Scores

Two measures of academic potential were used to assess the students: *Wechsler Adult Intelligence Scale-Revised* (WAIS-R) and the *Woodcock-Johnson Psychoeducational Battery Tests of Cognitive Ability (WJ)*. The data from these two measures is presented in Table 6 reveal two trends consistent with previously reported research: (a) the LD students demonstrate lower acadmeic potential than their peers, and (b) the *Wechsler* scale scores are higher than the scores obtained on the WJ. Since the WJ was specifically constructed to assess scholastic aptitude whereas the *Wechsler* scale was developed as a global measure of intelligence, the scores may reflect these differences in test purposes. LD students, score lower on the WJ compared to the *Wechsler* can be explained by the specific learning disability which lowers the student's aptitude or potential for academic achievement. However, a nonhandicapped student with average intellectual potential should be expected to demonstrate at least average scholastic aptitude. Therefore, the question arises whether or not the score differences obtained in these samples are the

Table 6

Standard Scores of Academic Potential

Group	LD			LA			RS			NR LD		
Test		SD	N	$\bar{X}$	SD	N	$\bar{X}$	SD	N	$\bar{X}$	SD	N
WJ	82.2	9.4	27	94.5	15.6	24	107.7	14.5	27	85.8	8.9	5
WAIS-R	93.1	10.1	13	93.2	15.6	6	113.2	9.3	9	91.3	10.9	6

results of sampling error. The impact of lower WJ scores would become apparent in the use of standard score discrepancy methods. A discrepancy is indicated by the difference between the student's expected achievement, as measured by an aptitude scale, and the actual achievement. If the WJ consistently yields lower scores than the Wechsler scale, fewer students would be identified as having a significant discrepancy when the WJ score is included as the measure of aptitude; that is, the calculated discrepancy scores would be smaller.

According to the pilot study data the random sample had a mean standard score one-half of a standard deviation above the population average. In contrast, the LD samples' means were at least one-half of a standard deviation below average on the Wechsler scale and one standard deviation below average on the WJ. The low-achiever sample scored between the LD and random samples on the WJ, and the same as the LD on the WAIS-R. The mean difference between the WAIS-R and the WJ was smallest for the LA group. Again, larger sample sizes are needed to confirm these data.

Comparison of Discrepancy Methods

Scores from the aptitude and achievement measures obtained on the LD, LA, and random samples of students were used in calculating discrepancy scores for four discrepancy methods: (a) Standard Score Comparison (after Hanna, Dyck, and Holen, 1979) (b) Regressed True Score Method (after Linn, 1982), (c) Regression Prediction Method (after Cone & Wilson, 1981; Thorndike, 1963), and (d) True Score Method (after Linn, 1982). However, an analysis of aptitude level and discrepancy methods was not completed because of the small sample sizes since the computational procedures of the discrepancy methods may favor the inclusion of students at some IQ levels (e.g., IQ scores above 110) while excluding students at other levels (e.g., IQ score less than 80). Until sample sizes of sufficient size are available for evaluating the discrepancy methods at different aptitude leves, we must rely on the use of theoretical comparisons as conducted in Appendix W in the paper "Comprehensive Analysis of Four Discrepancy Methods".

The pilot study data lend themselves to describing the distribution of discrepancy scores among the three sample populations. Specifically, the data allow an evaluation of the percent of the sample populations which had a discrepancy of a specific magnitude under each of the four discrepancy methods. For the purpose of this comparison, two incidence levels, 4% and 10% were chosen as magnitude criterion levels. The 10% incidence level is a more lenient cutoff than the 4% level; thus, more students should meet this criterion.

Table 7 lists the percent of the student sample who would qualify at a given incidence level for the four discrepancy methods. For example, the second column grouping refers to Method 2. In the achievement area of reading, for example, 27% of the currently identified LD population in the sample qualified at the 10% incidence level. In contrast, only 7% of the low-achiever sample met this criterion compared to 23% of the normal sample. Based on these data, one may conclude that a significant discrepancy occurs in the randomly selected sample of the community college almost as often as in the LD defined population. The following two observations should also be considered: (a) an achievement-aptitude discrepancy does not appear to be a discriminator between LD students and their normal peers, and (b) since LD students had significantly lower IQ than the normals, whose aptitude scores were average to above

Table 7

Percent of Students Meeting Discrepancy Criteria

Statistical Discrepancy Method*:	Method 1		Method 2		Method 3		Method 4	
Incidence level:	4%	10%	4%	10%	4%	10%	4%	10%
Achievement area: Reading								
Group: LD (N = 49)	6%	6%	10%	27%	22%	35%	12%	27%
Low Achiever (N = 30)	3%	3%	7%	7%	7%	7%	7%	7%
Normal (N = 36)	8%	8%	11%	23%	8%	14%	11%	19%
Achievement area: Written language								
Group: LD (N = 50)	12%	12%	18%	28%	32%	46%	18%	30%
Low Achiever (N = 30)	7%	7%	7%	17%	10%	23%	7%	17%
Normal (N = 36)	8%	9%	8%	33%	11%	17%	17%	33%
Achievement area: Mathematics								
Group: LD (N = 50)	4%	4%	8%	16%	14%	26%	8%	18%
Low Achiever (N = 50)	0%	0%	0%	3%	3%	10%	0%	3%
Normal (N = 35)	11%	11%	20%	23%	14%	23%	20%	23%

* Discrepancy methods

Method 1: Standard Score Comparison (after Hanna, Dyck, & Holen, 1979)
Method 2: Regressed True Score Method (after Linn, 1982)
Method 3: Regression Prediction Method (after Cone & Wilson, 1982)
Method 4: True Score Method (after Linn, 1982)

average, the normals were more likely to demonstrate a significant discrepancy. Essentially, this is due to the restricted ranges of the tests and normal variability in student performance. When comparing the percentages of the three subpopulations the reader is cautioned that the sample sizes are such that the addition of one or two subjects who meet the criteria can significantly increase the percentage values. Also, when evaluating the relative percentages consideration should be paid to the sample size on which they were based (listed in the left column for each group) and the three achievement areas.

The four discrepancy methods show a great deal of variability for the students who meet the criterion level. The method with the fewest number of students meeting the criterion levels was Method 1, the Standard Score Comparison Method. In contrast, Method 3, the Regression Prediction Method, consistently identified the most students. Method 2, the True Score Regression Method, and Method 4, the True Score Method, identified nearly the same percentages of students. The percentage differences across the four methods reflect computational differences between methods; that is, methods vary computationally in determining the discrepancy score, or the criterion score against which it is compared. Consequently, the table cannot be interpreted independently of some understanding of the mathematical calculations of the discrepancy methods. For example, the selection of a "preferred" method should not be based on the sole criterion of identifying the fewest students (e.g., Method 1) or the most students (e.g., Method 3) but should also take into consideration the assumptions on which the calculations are based.

Summary

This pilot study examined the aptitude, achievement, and discrepancy scores of LD students, Low-Achieving students, and Nonhandicapped students in a selected number of California community college learning disability programs. The primary objective of the study was to establish the groundwork for a proposed validation study of the discrepancy model and a description of current LD population.

The pilot study data were collected and 14 community colleges which were representative of geographic regions, urban, suburban, and rural populations, single- and multiple-campus colleges, and large, medium, and small student populations. A detailed set of written instructions for collecting the necessary aptitude and achievement data was presented to participating LD specialists.

The sample of LD students was chosen according to the following criteria: (a) the student must meet the state definition for LD, (b) the LD specialist's clinical judgment clearly suggested that the student was learning disabled, and (c) the student must have been in the CCC LD program less than one year or not have responded to the interventions of the LD program.

The low-achieving subjects were identified as those students who had been referred to the LD program as having possible learning disabilities and who did not meet the criteria for LD placement based on standard evaluation. Students were excluded from the LA group if their low achievement could be attributed to developmental disabilities, mental retardation, emotional disturbance, physical handicap, or sensory deficits. Selected subjects could best be considered by LD specialists as "low achievers".

The nonhandicapped students were those randomly selected students who by self-report had never attended or been referred to special education programs.

The final selection of participants from each of the three groups was based on a random selection procedure of students currently enrolled in or having completed eight credit hours. Students' aptitude and achievement test scores were taken from existing student records if the scores: (a) were no more than one year old, and (b) were judged by the LD specialists to be valid and reliable. The LD specialists tested students whose scores did not fit these two criteria.

While data were received from the participating LD specialists, the small sample sizes warranted only tentative conclusions. The aptitude scores of this sample showed two trends which were consistent with previously reported research: (a) the LD students demonstrated lower academic potential than their peers, and (b) the Wecshsler scale scores were higher than scores obtained on the Woodcock-Johnson Cognitive Ability Test.

Comparisons of the aptitude and achievement scores were made by using four methods for calculating discrepancy scores. The following observations were made: (a) an achievement-aptitude discrepancy does not appear to be a discriminating datum between LD students and their peers; (b) since LD students had significantly lower IQ scores than their normal peers, whose aptitude was average to above average, normal students tested were more likely to demonstrate a significant discrepancy; and (c) the LD students' mean aptitude score corresponded to their assessed achievement scores. That is, as a group, the LD students looked like low achievers.

SELECTION MODEL PILOT STUDY

Activity 1: Development of Selection Model pilot Survey instruments

Procedures. Three considerations guided the construction of the survey questions for the Selection Model pilot study. First was the need to focus as much as possible on behavior, that is, what the respondents do as opposed to what they think or believe. The Critical Incident Technique (Flanagan, 1954) was incorporated into the survey questionnaires to help accomplish this goal. Basically, the Critical Incident Technique is a method for collecting observed incidents which have special significance to the problem under study. Subjects are asked to identify significant incidents and to describe what happened based on the behaviors of the participants in the incident. The results are a description of actual behaviors that have some real effect on the problem.

A second concern was to design instruments that would collect the needed information without being so time consuming that respondents would wear out before finishing or hurry through the task without sufficiently considering their answers.

The third factor concerned the heterogeneity of the LD population and the equivocal indicators of a learning disability. Since previous research has not yet defined a set of distinct behaviors which consistently differentiate LD students from other low achievers, two courses of action were proposed. The first was to state behaviorally specific appropriate items from previous research. Hopefully, by stating specific distinct behaviors, LD subjects could be reliably differentiated from other low achievers. The second plan was to prepare items which sampled from behavioral domains which had previously received minimal attention.

Using these two methods, behavioral items could be researched within the context of the community college setting.

Several types of information were identified as being relevant to the selection pilot study: student performance factors, LD and regular classroom program demands, educational histories, LD program descriptions, and feasibility evaluations. In collecting at least part of these data, more than one desirable respondent could be identified. Thus, instruments were designed which would address all these factors in addition to soliciting information from college instructors and LD specialists as well as the SLD, USLD, and LA students. A description of the development of each instrument is presented below.

Program Demands Survey. A study of program demands recently completed at the KU-IRLD (Knowlton, 1982) included a comprehensive survey of program demands on secondary-level LD students across a number of factors: study skills, independent work habits, communicative skills, socialization and subject matter skills. As this instrument had already been field-tested as part of the Knowlton study and found to be satisfactory, it was selected as one of the selection pilot study instruments. For this purpose, it was modified slightly so that the respondents were asked to evaluate the degree of difficulty the subject (SLD, USLD, LA) had meeting each program demand. Several versions of this instrument were designed dependent on the group serving as respondents. For example, Form F for student respondents--SLD, USLD, LA--asked students to respond in terms of their own experiences. Form I for college instructor respondents required respondents to answer questions in terms of LD students in general rather than a specific LD student. Form H was intended for LD specialist respondents. For each LD student participating in the study, LD specialists responded based on their perception of the degree of difficulty each program demand presented. Copies of the Program Demands Surveys and accompanying instructions (FORMS F, H & I) are included in Appendix SS.

Survey Questions. The instrument entitled Survey Questions was designed to collect data on factors related to overall student performance with a subfocus on factors related to success within the community college program. The questions eventually selected for inclusion came from three sources: (a) criteria for distinguishing SLD students at the community college level elicited from LD specialists at the CAPED convention in Los Angeles October 23, 1982 (a listing of these criteria is included in Appendix TT); (b) LD adults' performance characteristics described in research reports published by the KU-IRLD; (c) characteristics which emerged in discussions of the Selection Pilot Study staff as being potentially useful for differentiating the targetted subgroups of community college students. These characteristics were employed as a safeguard against missing any significant characteristics through the other two sources.

Survey questions were written to collect information on the following factors: identification and attainment of community college goals, academic behaviors (e.g., attention, following directions, homework, comprehension difficulties), social interactions, self-help behaviors, student perceptions of factors accounting for their difficulties and successes within the regular community college classroom, English as a secondary language. The Critical Incident Technique was incorporated into this questionnaire by first asking students to think of times when (a) the teacher and (b) the student did something that significantly helped or limited the student's performance in the class, and then to describe what the teacher or the student did, and finally why this behavior had the effect it had.

As with the Program Demands survey, three versions of the Survey Questions were needed: Form G for student respondents (SLD, USLD, LA); Form H for college instructor respondents; and Form E for LD specialist respondents (Appendix UU). As in the Program Demand surveys, students responded based on their own experiences, college instructors answered the questions for LD in general, and LD specialists answered based on their perception of each LD subject in the study.

ESL Questions. Because of immediate concerns related to servicing LD students who are not native speakers of English, a separate section of the survey questionnaire was devoted to English-as-a-second-language factors. These questions were asked only of students who were not native English-speaking students (Appendix VV).

Student Data Sheets. In addition to the student information described above, information specific to each student's educational history and academic achievement was collected using the Student Data Sheet (Form C).

The students served as informants for questions related to educational history and were also asked to rank order areas of academic achievement to determining which of them had presented the most difficulty throughout their education. The LD specialists completed the section of the data sheet asking for specific scores on measures of intellectual aptitude and academic achievement. The student data sheet is included in Appendix WW.

Program Description Instrument. The initial intent behind the pilot study was to collect detailed information about the LD program of each participating community college. Such descriptions would be crucial for judging student/ treatment interactions and other factors related to student performance in the LD program. A comprehensive survey was designed to meet this need, However, as the pilot study evolved, it became obvious that the LD Program Description Questionnaire was too time consuming to use given the time constraints of the CC staff and the sutdy. As a result, this element was removed from the pilot study but retained for possible use in the later validation study. In its place, a shortened questionnaire was devised which solicited information about student referrals, placements, and withdrawals from the LD program. This information was judged useful for evaluating the impact of the identification procedures to result from the discrepancy and selection pilot studies. The LD Program Description (Form L) is included in Appendix XX.

Formative Evaluation Sheet. As described in the discrepancy pilot study, input was sought from the LD specialists who implemented the selection pilot study to help evaluate the feasibility of the procedures if they were included in the recommended set of identification procedures. This input was solicited using the Formative Evaluation Sheet (Form D). Form D is included in the Appendix YY.

Activity 2: Preparation of Materials for Data Collection

Procedures. Keeping track of the data collection instruments represented a major challenge to implementers of the selection pilot. As in the descrepancy pilot study, three techniques were used to facilitate the management of the selection pilot. First, a detailed step-by-step description of every component of the study was prepared (Appendix ZZ). This description included 14 steps ranging from the initial examination of the research materials packet to mailing the completed materials at the end of the data collection. Within the outline, all materials were clearly identified by name and form letter. Second, all materials were color-coded to indicate the specific responsibilities of the various participants. The following color-code was used:

Respondent	Color code
SLD	blue
USLD	buff
LA	pink
College Instructor	yellow
LD Specialist	green

Third, a checklist of tasks (FORM M) (Appendix AAA) was designed to help the LD specialists keep track of the steps to be completed. The LD specialists were directed to record starting and completion dates for each task so that they could see at a glance what had been done, what was in progress, and what was yet to be done. This checklist also provided an estimate of how much time was spent on any one component of the study as well as total time. This informaition would be useful in evaluating the feasibility of the various selection model components in the identification procedures which will ultimately be developed.

The survey materials were assembled into packets which were later distributed to the participating LD specialists. The material packets were organized sequentially, that is, in the order in which each material would be used. Figure 2 contains a listing of the materials provided for the pilot study.

Figure 2

Contents of Selection Model Data Collection Packets

A. Consent Forms for students (FORM B)

B. Stamped envelopes for returning consent forms

C. Student Surveys: 12 copies of each

  Part I: Program Demands (FORM F)

  Part 2: Survey Questions (FORM G)

  Student Data Sheet (FORM C)

  Cover Letter for consent forms (FORM N)

D. College Instructors' Survey: 4 copies

  Cover Letter (FORM K)

  Part 1: Program Demands (FORM I)

  Part 2: Survey Questions (FORM J)

E. LD Specialist Surveys: 4 copies of each

  Part 1: Program Demands (FORM H)

  Part 2: Survey Questions (FORM E)

  Part 3: Student Data Sheets (FORM C)

F. Manila envelope (size 6½ X 9'): 8 copies

G. Manila envelope (size 11½ X 14½) addressed to IRLD, University of Kansas

H. List of random numbers

I. Formative Evaluation Sheet (FORM D)

J. Record of Student Refusals (FORM A)

K. Program Description (FORM L)

L. Checklist of Tasks (FORM M)

Activity 3: Selection and Training of the Selection Pilot Task Force

Procedures. As described in the discrepancy pilot study, a pilot study task force comprised of 25 LD specialists in California's community colleges was created to implement the two pilot studies. These volunteers were selected in order to obtain a representative sample of community college programs throughout the diverse California community college system. Fifteen of the 25 Task Force members volunteered to implement the selection pilot. Four of the 15 also collected data for the discrepancy pilot study. A list of the selection pilot study task force is included in Appendix BBB.

Training of the selection task force members was conducted on February 9th and 10th at DeAnza College for Northern California participants; Cypress College was the training site for Southern California volunteers. As described for the discrepancy pilot study, the training sessions for the two pilot studies were held back-to-back on a single day in each training site to reduce the travel demands on participants. Training for the selection pilot study involved the distribution of the materials packet, a review of the detailed outline of tasks, and a question-answer session to clear up any immediate concerns and confusions. Steering Committee members followed up through phone calls or visits to each of the participating campuses by Steering Committee members to review the tasks and to problem solve with the participants.

Activity 4: Selection of Research Subjects

Procedures: Three groups of respondents served as subjects for the selection pilot study: (a) community college students; (b) community college instructors; (c) LD specialists.

The community college students employed in the study consisted of three subgroups: (a) LD students who were successful in meeting their goals for attending community college; (b) LD students who were unsuccessful in meeting their goals for attending community college; and (c) low-achieving students. A total of two successful LD (SLD), two unsuccessful LD (USLD) and four low achievers (LA) were identified at each of the 15 community colleges participating in the study. Thus 30 SLD, 30 USLD, and 60 LA students were projected as participating in the study.

Two subgroups of Community College instructors were included in the study: (a) instructors identified by the LD students as having taught a class in which they experienced success; (b) instructors identified by the LD students as having taught a class in which they felt unsuccessful. Two "successful" and two "unsuccessful" instructors were identified at each participating community college for a total of 30 "successful" and 30 "unsuccessful" college class instructors.

Finally, the LD specialists comprised the 15 LD specialists who had volunteered to implement the selection model pilot study at their own community colleges.

The following criteria and procedures were used to generate a pool of potential LD, LA, and college instructor participants and to make final subject selections.

A. Criteria for LD Group

1. The students must have left the LD program in 1982. Reasons for "learning the program" included graduation, dropping out, mainstreamed into regular classes, etc., but not moving to another location.

2. The students must qualify as LD based on the following definition: A specific learning disability refers to disorders in which an individual exhibits a significant/severe discrepancy between the current level of developed intellectual abilities and adademic performance despite regular instruction and educational opportunity, as currently measured by professionally recognized diagnostic procedures. Academic performance refers to achievement in the following areas: listening comprehension, oral expression, written expression, basic reading skills, reading comprehension, mathematical calculation, and mathematical reasoning. Specific learning disabilities are often due to constitutional, genetic and/or neurological factors, and are not primarily due to: Visual (blind) or auditory (deaf) sensory deficits, motor handicaps, severe emotional disturbance, environmental or economic disadvantage, cultural/language difference or mental retardation.
3. The LD specialist should have no doubt that the students meet criteria 1 and 2. The names of all students meeting these criteria were listed for use in final subject selection procedures.

B. Criteria for Low-Achieving Group

1. The students were referred to the LD program as having possible learning disabilities;
2. As a result of the LD program's standard evaluation procedures (preferably a comprehensive evaluation), it was determined that the students did not meet the criteria of the LD program;

3. The students' low achievement was not attributable to developmental disabilities, mental retardation, emotional disturbance, physical handicap, or sensory deficits.
4. The students are best considered as "low achievers".
5. The LD specialist should have no doubt that the students meet criteria 1, 2, 3, 4 and 5.

The names of all students meeting the low-achiever criteria were listed for use in the final subject selection.

From the lists of LD and LA students who met the minimum criteria for inclusion in the study, actual participants were selected. The following sections explain the procedures used to make the final subject selections.

C. Identifying "Best Fit" LD Students

The usefulness of the data collected on the LD subjects depended directly on how well the actual LD participants fit the criteria described above for the LD group. To increase the likelihood that the LD participants were indeed LD, the LD specialists used the following procedures to ensure that the selected subjects were those who best fit the LD criteria described in A above.

1. From the roster of LD students who met the LD criteria, the LD specialists ordered students according to those who "best fit" the LD definition. In this ranking, the first student should be the one who, in the LD specialists' judgment, "best fit" the LD definition.
2. The LD specialist continued ordering students until they had a list containing 25% of the original list. For

example, if 50 LD students had left the program in 1982 and met the other criteria, 13 students (25% of 50) would be listed on the "best fit" list.

D. Identifying Successful and Unsuccessful LD Students

From the list of LD students rank-ordered for best fit, two successful and two unsuccessful students were identified. These four students were to serve as the LD subjects for the selection pilot study. The following procedures were used to identify these four students.

1. Successful LD Students. Two students from the list of "best fit" LD students, who in the LD teacher's judgment were successful based on the following definition: A successful student is one who has met his/her goals for attending the community college program.
2. Unsuccessful LD Students. Two students from the list of "best-fit" LD students who in the judgment of the LD specialist were most unsuccessful in meeting their goals for attending the community college program.

E. Identifying Low-Achieving Students. Low-achieving students were randomly selected using a random number table and the following procedures.

1. Each name on the roster of LA students generated in B above was assigned a number according to its position on the list, that is, the first name became 1, the second name 2, and so on. These numbers were used to match names with random numbers selected from a random number list provided in the materials packet.

2. A starting place on the random numbers list was marked in red ink. ,The following example was provided to direct the LD specialist in matching names and random numbers.

Suppose that your starging number is 104. Follow these steps.

A. Does my roster have 104 names?

B. If yes, locate the name in position number 104.

C. If no, to go the next number on the list which would correspond to a name. In moving through the list you can go in any direction you wish (up, down, left, right, diagonally, etc.) but do not change direction once you have started. Your list of names may be fewer than the numbers on the random number list (which is 50,000). In this case, you must consider random number values that are smaller. For example, if your list of names only includes 98 names, you would only consider the last two digits for each random number. Similarly, if your list included nine names, you would only consider the last digit from each random number.

D. Continue matching names with numbers until you have two students who meet the qualifications for the group and agree to participate.

E. As you select students it is wise to include several alternates in case that those initially chosen are not able to participate.

Once the final set of student participants was generated, the LD specialists contacted the selected students and obtained their consent to participate. The teachers received an outline of points to cover in this initial contact. The main points included a description of: (a) the project as one designed to study the factors that lead to student success in the community college; (b) the student's participation in a phone or in person interview at the community college in which questions would be asked about his/her community college exepriences; (c) the expected time needed (less than two hours); (d) assurance of complete confidentiality of the student's responses and participation.

When contacting the LD students, the LD specialist was also to ensure that the students view of their college experiences as successful or unsuccessful corroborated the LD specialist's classification of the students. This check was accomplished through a screening question: "When you consider your experiences at the community college, did you meet your goals for attending?" If the students answer did not match the LD specialist's classification, teachers thanked them adding that they would not be needed after all, and selected the next name on the list as an alternate. If the names on the initial "BEST FIT" LD list were exhausted before two successful and two unsuccessful students agreed to participate, no other students were contacted as alternates.

Once students agreed to participate, they received a copy of the surveys (Forms C, F, and G) so that they would have an opportunity to consider their answers more thoroughly. Also included in the mailing was a cover letter (FORM B) whitch described the study and the students' role, and a Consent Form (Form B) (Appendix PP) to be signed and returned in a pre-addressed, stamped envelope.

If a student refused to participate, the LD specialist attempted to determine why and recorded the reasons on Form A (Appendix CCC).

Activity 5: Selection of College Instructors

Procedures. Selection of college instructors to participate in the selection pilot study was based on the LD students' responses to a survey question asking them to name a college instructor and a class in which they felt "successful" and one in which they felt "unsuccessful". Thus, the college instructors were selected after the LD student surveys had been completed. The following set of directions was provided to the LD specialists to help them select two instructors who taught classes in which LD students felt successful and two instructors who taught classes in which LD students felt unsuccessful:

A. The students were asked to name an instructor and a college class in which they felt "successful" and "unsuccessful". Only the LD students' responses to those questons on the surveys are of interest.

B. If the LD student names a faculty member who is no longer at the community college, follow up his/her response by asking for the name of someone else and a course which came closest to meeting the criteria. This second name is for your information only; it does not change the student's responses.

C. After the four LD students have been surveyed, prepare two lists of names. On the first, list the four faculty names associated wth a "successful" class; list the four faculty associated with an "unsuccessful" class on the second list.

D. For the next step you will again need your table of random numbers.

E. Use only one digit from each number, the digit with the "ones" place value.

F. Using the "successful" list of names match the first random number with a name in the corresponding position on the list. Select two different names in this manner.

G. The resulting names signify the faculty members you are to ask to participate in the project. Repeat the selection steps with the "unsuccessful" list of names.

H. If a faculty member declines to participate, select another faculty member, using the table of random numbers.

The Checklist of Tasks (Form M) outlining the subject selection procedures provides a summary of the sequence of steps used to identify subjects for the selection pilot study. Form M is located in Appendix DDD.

Activity 6: Collection of Survey Data

Procedures. The data collection began immediately after the training session held on February 9th and 10th. April 15th, 1983 was targeted as the completion date.

Four sources were used for data collection within each of the 15 sites:

1. Students--4 SLD, 4 USLD, 4 LA
2. College instructors--2 successful; 2 unsuccessful
3. LD specialists
4. Student records

Student information was collected through telephone or in-person interviews conducted by the LD specialist whereas the college instructors and LD specialists were asked to complete questionnaires independently. The LD specialists were responsible for collecting student data from the student's records.

A sequence of procedures was recommended to facilitate the LD specialists' data collection efforts; however, the order of the steps could be altered if, the suggested data collection sequence was as follows:

1. Mail or give the Cover Letter (Form N), Consent Form (Form B), Student Data Sheet (Form C), Program Demands (Form F), Survey Questions (Form G) to student participants.

2. Complete the LD Specialist's Survey on each LD student--Program Demands (Form H), Survey Questions (Form E), and Student Data Sheet (Form C). This task must be completed before the student interviews are conducted.

3. Assign a group code (Item 3a on the Student Data Sheet) using the following scale: Successful LD = 1, Unsuccessful LD = 2, LA = 3.

4. Complete all LD student interviews: Program Demands (Form F), Survey Questions (Form G), Student Data Sheets (Form C).

5. Identify and survey college instructors

   a. Select regular class instructors and obtain their consent to participate.

   b. Assign each faculty member a unique code, for example, initials, a 5-digit random number, or anything that uniquely identified the person give each faculty participant a copy of the cover letter (Form K), program demands (Form I), and Survey Questions (Form J). Ask that Forms I and J be returned within one week.

   c. Check with the instructor in a few days to answer any questions.

   d. When the instructor's forms are returned, record the following information in the upper right-hand corner:

1. Successful--1; Unsuccessful--2
2. Course content, for example, English, History, PE, Math, etc.

6. Conduct interviews with LA students at any time since you will not complete survey forms on this group of subjects.

7. Complete the Program Description (Form L) at any time though part of it requires information generated from previous step.

As previously described, a Checklist of Tasks (Form M) was provided to help the LD specialist keep a running progress check and to collect information on how much time was required to complete each step of the pilot study. This checklist is located in Appendix DDD.

Activity 7: Collection of Formative Evaluation Data

Procedures. As described for the discrepancy pilot study, LD specialists were asked to evaluate the procedures and instructions provided for them in implementing the selection pilot study. This input was to be used to revise the selection procedures prior to initiating the comprehensive validation study scheduled for the following year and to formulate the LD identification procedures intended to be the final product of the research project.

Each selection pilot materials packet contained a Formative Evaluation Sheet (Form D in Appendix RR). For each step of the pilot, LD specialists were asked to evaluate the clarity of the instructions and the effectiveness of the procedures as well as to provide other comments and suggestions. Completion of the Formative Evaluation Sheet represented the final data collection procedure for the selection pilot. All research materials were returned to the University of Kansas Institute for Research in Learning Disabilities in a pre-addressed, stamped, manila envelope.

Data Analysis

The collected data were entered in a Honeywell 66/60 mainframe computer for statistical analysis. However, before the analysis, the organization of the data was adjusted. For example, several single questions on the survey were treated as a collection of items. Thus, question 11 on the Student Survey was treated as eight independent items--one for the primary area of difficulty and the remaining seven items for other difficulty areas. Moreover, the response choices for questions 42, 43 and 46 on the Student Survey were also considered as independent items. This arrangement resulted in a total of 117 items for analysis.

The data arising out of questions 4 to 11, 43, 44, and 46 on the Student Survey were treated as nominal level data on the measurement scale, whereas the remaining data were treated as continuous variables. Respondents were categorized into three groups: (a) successful LD, (b) unsuccessful LD, and (c) low achievers. One-way analysis of variance was performed to identify group differences in survey responses. Nominal data were analyzed by the Chi-square statistic ($\underline{X}^2$). All observed differences were tested for significance at the .05 level. Open-ended questions were excluded from this statistical analysis. Missing values were identified and were not included in the analysis.

The data analysis was completed in four steps. At step 1, successful LD students' responses were compared to those of unsuccessful LD students. The analysis was performed on 202 items 85 of which were from the survey completed by LD teachers.

In step 2, all LD students were compared with low-achieving students. Since the instruments for low-achieving students did not include a section to be completed by LD teachers, the Step 2 analysis was restricted to the student questionnaires, with a total of 117 items.

At step 3, three independent groups representing LD-successful, LD-unsuccessful, and low achievers were compared with each other on 117 items. The student-Newman-Keuls test was used as post hoc comparison on items that were judged statistically significant.

At step 4, the data obtained from regular classroom teachers were analyzed. Prior to the analysis these data were also divided into (a) successful teachers and (b) unsuccessful teachers. A total of 85 items were processed.

Mean and standard deviations were also computed on the variables under study. The Statistical Package for Social Sciences (Hull & Nie, 1981) was used as a computing software.

Participating Campuses

LD specialists from 15 campuses had initially agreed to participate, however, eventually the KU-IRLD staff received data from only nine campuses: Bruce Sturm at DeAnza College, Gill Hoanzl at El Camino College, Cheryl Dunn-Hoanzl at Santa Ana College, Kirsten Sondergaard-Colvey at Crafton Hills College, Pamela McCarter at Southwestern College, Inge Pelzer at Chaffey College, Betty Creson at Ventura College, Dinad Durham at Glendale College, and Jan Shapiro at Santa Barbara College. Explanations offered from other campuses for failure to complete and return the materials included: the tasks were more involved than expected, the campus' financial crisis had led to cutbacks in staff, and the task was too complex.

Sample Sizes

Since the amount of materials returned from each LD specialist varied, the sample sizes for each group were less than expected. If the total data had been returned, the sample sizes would have consisted of 60 students each in the successful LD group, the unsuccessful LD group, and the low-achieving group. However, the actual number of cases were as follows: (a) 18 students in the successful LD group; (b) 16 students in the unsuccessful LD group; and (c) 15 students in the low-achiever group.

Given the small sample sizes and the comparatively high number of variables in the analysis, the results must be considered tentative. In particular, it is impossible to estimate population parameters, such as for the learning disabled group, when the sample sizes are small and when the group scores vary extensively. Unfortunately, both of these conditions were present in the pilot study.

RESULTS

Statistically Significant Items Which Distinguished LD from Other LA Students

Within the pool of 117 items analyzed from the selection survey questions, a total of 14 items reached statistical significance in differentiating LD's from LA's. Each of these 14 items will be discussed separately. The small sample sizes, especially for the LA group, must be kept in mind when evaluating the generality of these results.

Student sex. Table 8 presents the percentage of males and females within the LD and LA groups. Of a total of 33 LD students, males outnumbered females by a ratio of nearly 2 to 1 (corrected $\underline{X}^2 = 4.26$, df = 1, $\underline{p} = .03$). In contrast, females outnumbered males by a ratio of nearly 3 to 1 in the LA group.

TABLE 8

Gender Of Participating Students

Student Group		Sex		
		1 = Male	2 = Female	Total
LD	n	21	12	33
	%	63.6%	36.4%	
LA	n	4	11	15
	%	26.7%	73.3%	

$\underline{X}^2 = 4.26$

Probability value = .0389

As higher proportions of males than females have frequently been observed in LD programs at both elementary and secondary grade levels, it is not surprising to find a similar situation at the college level. Despite these consistent findings, the reasons why males are more often placed in LD programs than are females have not been thoroughly investigated. Sex itself is unlikely to be the determining placement factor. Instead it may be an artifact associated with other factors such as severity, adaptability, resourcefulness, attitude, emotinal stability or socialability. Consequently it would be inappropriate to include sex as a determining factor in deciding eligibility for LD service; that is, if a male and a female student presented equivalent profiles on other selection criteria, it would be inappropriate to give preferential consideration for the LD program to the male student by virtue of his maleness.

Number of schools attended. The number in junior- and senior-high schools attended by LD and LA students is presented in Tables 9 & 10. Although statistically significant Chi-squares support the conclusion that overall LA students changed schools more often than did LD students

TABLE 9

Number of Attendance Centers: Junior-High Schools

	Student Group	Number of Junior-High Schools				Total
		0	1	2	3	
$\underline{x}^2$ = 15.09	LD N %	0 0%	25 75.8%	8 24.2%	0 0%	33
Probability value =.0017	LA N %	2 13.3%	7 46.7%	2 13.3%	4 26.7%	15

TABLE 10

Number of Attendance Centers: Senior-High School

	Student Group	Number of High Schools						Total
		0	1	2	3	4	5	
$\underline{X}^2$ = 12.40	LD N %	0 0%	25 75.8%	1 3%	5 15.2%	2 6.1%	0 0%	33
Probability value =.0296	LD N %	2 13.3%	7 46.7%	3 20%	2 13.3%	0 0%	1 6.7%	15

in junior-high ($\underline{X}^2$ = 15.0, df = 3 p = .001) and senior-high school ($\underline{X}^2$ = 2.40, df = 5, p = 029), the small sample sizes may be biasing the results. These findings must be confirmed with larger samples sizes as number of school changes is often used by LD practitioners to operationalize the exclusionary component in the definition of learning disabilities which relates to adequate educational opportunity. If an academically deficient student presents an educational history which includes numerous school changes, the latter may be considered sufficient evidence against an LD diagnosis.

Primary area of academic difficulty. Table 11 presents the findings related to LD/LA self-reports of academic areas which cause them the most difficulty. Reading recognition was the only academic category that attained statistical significance in differentiating the LD from low achievers. In this respect, LD students were more likely to rate reading recognition as an area of difficulty than were LA. This finding may be particularly useful in the selection model as that valid and reliable tests are available for confirming reading recognition difficulties.

TABLE 11

Reading Recognition As Primary Area of Academic Difficulty

	Student Group		Primary Area of Difficulty		Total
			No	Yes	
Reading Recognition	LD	n	13	20	33
Corrected $\underline{X}^2$ = 4.57		%	39.4%	60.6%	
Probability value = .32	LA	n	11	3	14
		%	78.6%	21.4%	

Academic achievement. Differences between LD and LA scores on both the Woodcock-Johnson reading cluster scores and written language cluster scores attained statistical significance. Table 12 presents the mean standard scores and standard deviations for each group. LD students scored at a lower level than did low achievers on both the reading and the written language clusters. These results gain support by the limited overlap of the 95% confidence interval for LD and LA on the reading cluster means and the absence of an overlap of confidence intervals for the written language cluster means. Also, these findings are consistent with those obtained for LD and LA students on the Woodcock-Johnson Achievement Battery reported for the discrepancy model pilot study.

TABLE 12

Academic Achievement Of Learning Disabled And Low-Achieving Students

Woodcock-Johnson Reading Cluster Score

F - Value = 5.565		LD	LA
Probability value = .28	X	79.13	94.25
	SD	15.47	13.26
	N	16	8
	95% conf. interval	70.88-87.37	83.16-105.24

Woodcock-Johnson Written Expression Cluster Score

F - Value = 6.621		LD	LA
Probability value = .018	X	78.6	93.0
	SD	14.9	6.8
	N	15	8
	95% conf. interval	70.35-86.85	87.31-98.6924

Academic behaviors. LD and LA self-ratings differed at a statistically significant level on eight of the selection items related to academic behaviors. Means, standard deviations, sample sizes, and 95% confidence intervals are reported in Table 13.

LD's reported encountering greater difficulty than low achievers in

BEST COPY AVAILABLE

TABLE 13

Academic Related Behaviors

Distinguishing Learning Disabled and Low Achieving Students

Test-Taking Skills

		LD	LA
F - ratio = 4.135	$\bar{X}$	2.94	4.08
Probability value = .048	SD	1.72	1.66
	N	32	13
	95% conf. interval	2.32-3.56	3.08-5.08

Working Beyond Expectation*

		LD	LA
F - ratio = 6.636	$\bar{X}$	4.74	5.87
Probability value = .013	SD	1.66	0.52
	N	34	15
	95% conf. interval	4.16-5.31	5.58-6.15

Participating in Discussion*

		LD	LA
F - ratio = 7.403	$\bar{X}$	4.00	5.40
Probability value = = .008	SD	1.82	1.18
	N	33	15
	95% conf. interval	3.35-4.65	4.74-6.06

Reading Skills*

		LD	LA
F - ratio = 9.424	$\bar{X}$	3.38	5.08
Probability value = .004	SD	1.79	1.38
	N	34	13
	95% conf. interval	2.76-4.01	4.24-5.91

Spelling Skills*

		LD	LA
F - ratio = 9.682	$\bar{X}$	2.53	4.29
Probability value = .003	SD	1.76	1.82
	N	34	14
	95% conf. interval	1.91-3.14	3.24-5.33

Learning from Printed Material*

		LD	LA
F - ratio = 4.211	$\bar{X}$	4.32	5.29
Probability value = .046	SD	1.61	1.07
	N	34	14
	95% conf. interval	3.76-4.89	4.67-5.90

How Frequently Did You Seek Tutorial Assistance to Make Good Grades?**

		LD	LA
F - ratio = 10.738	$\bar{X}$	2.85	4.20
Probability value = .002	SD	1.28	1.42
	N	34	15
	95% conf. interval	2.41-3.30	3.41-4.99

How Frequently Did You Seek Help from a Tutor?**

		LD	LA
F - ratio = 14.420	$\bar{X}$	2.85	4.67
Probability value = .001	SD	1.40	1.84
	N	34	15
	95% conf. interval	2.37-3.34	3.65-5.68

How much help did You Need to Complete a Typical Academic Task?***

		LD	LA
F - ratio = 6.134	$\bar{X}$	3.42	2.43
Probability value = .017	SD	1.23	1.34
	N	33	14
	95% conf. interval	2.99-3.86	1.65-3.20

*Responses were recorded on a scale where:

1 = Very difficult
2 = Quite difficult
3 = Somewhat difficult
4 = Slightly difficult
5 = Difficult
6 = Not difficult

**Responses were recorded on a scale where:

1 = Always
2 = Usually
3 = Sometimes
4 = Seldom
5 = Almost never
6 = Never

***Responses were recorded on a scale where:

1 = Less than others
2 = Almost the same
3 = Slightly more than others
4 = More than others
5 = Almost double
6 = More than double

test-taking skills, working beyond expectations, participating in discussions, reading and spelling skills, and learning from printed materials. No overlap between the 95% confidence intervals was found for working beyond expectation, participating in discussions, and reading and spelling skills. This finding adds considerable weight to these four items in that it is unlikely that the true group means overlap.

Test-taking skills and spelling skills represented the areas of greatest difficulty for the LD group. Thus, the mean rating for these two items fell between "quite difficult" and "somewhat difficult". These two areas also received ratings of comparatively greater difficulty for LA students. However, in contrast to the LD subjects, the LA's reported having only "slight difficulty" with these skills.

Two items dealt with the frequency with which students sought tutorial assistance. On both items, LD's reported seeking such assistance much more frequently than did LA's. Again, the 95% confidence interval for the two groups did not overlap. Overall, the LD's true mean fell between "usually" and "sometimes seek assistance" whereas the LA's true mean ranged between "seldom" and "almost never seek assistance".

The above findings may have major significance for differentiating those students who in the past have been considered LD as opposed to LA. Students who voluntarily seek help to improve are often those who will receive help whereas students who do not show such a self-help attitude are unlikely to have help forced upon them. From a treatment efficacy standpoint, it is beneficial to select students for LD service who have demonstrated an interest in receiving such service.

The last item on which LD and LA demonstrated statistically significant differences dealt with the amount of help the student reported needing to complete an academic task. Here the LD's reported needing between "slightly more" and "more" than others whereas LA subjects indicated that their needs in this area fell somewhere between "almost the same as others" and "slightly more than others". Despite this statistically significant finding ($p = .017$) considerable overlap was found between the confidence intervals for the two groups on this time, suggesting that the true means may not differ as much as statistical significant testing would indicate.

Summary

The selection model pilot study was developed to identify qualitative factors on which learning disabled students can be distinguished from other low achievers. Since previous research has failed to identify characteristics unique to the learning disabled population, this study employed multiple survey methods of several respondent groups. Respondents included three different groups of students, community college instructors, and LD specialists. Four surveys and interview forms were used to obtain information about (a) the type of academic demands placed on students; (b) the factors that influence successful or unsuccessful performance within the community college; (c) the assessed aptitude and achievement test characteristics of the students; and (d) the unique problems of ESL-NES students in LD programs. Data were received from 9 of 15 campuses which had agreed to participate. This response resulted in a sample size of 49 students (15 LA and 34 LD).

For the populations studied, the LD and LA students were alike. However, differences were found on 12 variables in five performance areas:

1. ACADEMIC PERFORMANCE: LAs were comparatively better than LD on Woodcock-Johnson Reading and Written Expression cluster scores and ability to participate in discussions, work independently, learn from printed material, and to learn reading and spelling, and reading recognition skills.

2. SELF-HELP: LD sought tutorial assistance more often than did LA.

3. SEX: Disproportionately fewer females were classified as LD then males.

4. EXCLUSIONARY FACTOR: LAs changed junior-high and high-schools more frequently than did LD. This finding may reduce the possibility of an LA being considered LD.

5. SOCIAL ACTIVITY: LAs were more active in social organizations than LD students.

In conclusion, the results of this pilot study parallel the findings in the discrepancy model study, that is, few reliable indicators can be found to distinguish the LD population from other low achievers. This minimal success at identifying a criterion group of LD students suggests that the definition of learning disabilities should be changed or that alternative eligibility criteria are needed.

SUMMARY

Results of recent studies undertaken by the California Department of Finance and the Office of the Chancellor of California's Community Colleges have demonstrated considerable variability in the procedures and criteria used to identify learning disabled (LD) students within the California Community College System. The documentation of a lack of consistency in LD identification procedures has resulted in growing concern that statewide inequities may exist in: (a) the categorical aid (excess costs) provided by the state to community college LD programs; and (b) the availability of service to students as the eligibility criteria of one LD program can be more or less exclusive than those of other LD programs. Consequently, the Chancellor has asked that procedures be developed leading to minimum eligibility standards for identifying LD adults in the community college programs.

It is not surprising that LD programs in California community colleges lack consistent eligibility criteria. Current federal policy bases the identification of a learning disability on the existence of a significant discrepancy between a student's measured intellectual ability and academic achievement in the absence of any other primary handicapping conditions or disadvantage. According to the Federal Register (December 29, 1977), "when a severe discrepancy exists which cannot be explained by the presence of other known factors that lead to such a discrepancy, the cause is believed to be a specific learning disability" (p. 65085). However, attempts at using PL 94-142 guidelines to identify LD students have proven problematic. The concept of ability/achievement discrepancy, though crucial to the definition of learning disabilities, is not easily operationalized, particularly at the community college level. Furthermore, the federal regulations do not adequately address the difficulty of differentiating LD from other low-achieving students.

More specifically, the following questions remain unsolved: (a) how large of an academic discrepancy is necessary to determine a learning disability; (b) which tests are valid and reliable for use with adult-aged populations; (c) what abilities and skills should be measured; and (d) what exclusions should apply. An additional concern is that an academic discrepancy alone does not adequately tap the true nature of a learning disability, i.e., a handicap in the ability to learn, not just a measured difference between what an individual has actually learned and what he/she would be expected to have learned. All of these questions make the inconsistencies in California's community college LD identification procedures inevitable. Indeed, similar inconsistencies abound nationwide in LD programs for all age groups. However, although the need for a comprehensive study of this complex problem is obvious, to date the California Community College System is the only major educational system to systematically address each of the problems outlined above.

Statement of Problems

The development of minimum eligibility standards for identifying LD adults is surrounded by a broad array of very complex issues. Unless each of these is dealt with systematically, validated identification procedures for LD populations are not likely to be generated. The major problems that have been addressed as part of the present project will be delineated in the following paragraphs.

1. Differentiating attributes have not been specified. At the heart of the LD identification problem is the ability to specify those learner

attributes or characteristics that best differentiate the person with learning disabilities from individuals who are doing poorly in academic settings because of general low achievement. To date, researchers have been unable to identify any single attribute that is sufficient (by itself) to indicate a learning disability, or to specify a learning disability which is so debilitating as to warrant specialized educational opportunities. In short, existing attributes are, at best, equivocal indicators of a specific learning disability. This problem is exacerbated by the paucity of empirical data on the attributes which reliably differentiate the learning disabled from the larger group of other low achievers. Against this background, the LD specialist is confronted with the problem of deciding whether or not a student is learning disabled. Not surprisingly, therefore, this confusion has resulted in inconsistent identification practices and numerous selection errors.

2. Inappropriate assessment instruments are used in the identification process. Accurate identification of an LD population requires the use of appropriate psychometric testing instruments to quantify the decision-making process. A recent study sponsored by the Chancellor's Office (Ostertag & Baker, 1982) reported on the psychometric instruments most frequently used for identification and assessment purposes. The findings of this investigation revealed two basic problems inherent in current identification practices. First, several test instruments currently used suffer from a lack of sufficient technical adequacy (e.g., lack of age-appropriate norms) for a community college population. Consequently, the application of such instruments will not yield valid data. Second, Baker and Ostertag (1982) reported great inconsistency in the use of selected instruments. That is, some colleges use a given instrument for identification purposes, whereas other colleges use the same test for making diagnostic inferences (e.g., specifying particular deficits in a student's math functioning).

3. Sufficiently powerful (sophisticated) model(s) for measuring achievement discrepancies have not been specified. The statistical concept of a discrepancy between a student's actual achievement (as measured by a standardized achievement test) and his/her predicted achievement (as measured by a standardized aptitude test) is one of the most commonly cited characteristics of a specific learning disability. In general, LD students are seen as achieving below their predicted level of achievement. Of the numerous problems related to the operationalization of the discrepancy concept, the following are among the most critical: (a) how should the actual discrepancy score be calculated (e.g., raw score method, residual score method, true score method, regression method, etc.)? and (b) how should the discrepancy score be interpreted (e.g., statistical significance, educational significance, etc.)?

The more popular, traditional methods to measuring and interpreting discrepancy scores are fraught with naive and inaccurate assumptions. The utility of such procedures lies in the simplicity with which they can be calculated. In part, the reliance on the simpler methods may stem from a lack of understanding the more technically adequate methods. Most importantly, however, any method for calculating a discrepancy score must demonstrate sufficient technical sophistication to evaluate the influence of such statistical measurement factors as: measurement error, regression effects, and increased variability of achievement scores with increasing age.

The effects of these factors must be evaluated through a study of various discrepancy methods. Unless the methods are evaluated, numerous selection errors will occur, such as the exclusion of a disproportionately large group of students whose aptitude scores are below or above average.

Unfortunately, no empirical evaluation of these alterntive discrepancy methods has been undertaken.

4. Other criteria must be specified in addition to aptitude and achievement measures. While the discrepancy between a student's assessed achievement level and predicted achievement level constitutes a commonly agreed upon criterion for specifying a learning disability, it is not sufficient by itself to determine the presence of a learning disability. Unless other factors are carefully considered, an identification procedure responsive to the needs of the LD population will not result. Thus, the following problems arise from relying solely upon a discrepancy criterion. First, even after the discrepancy scores of a group of students have been calculated and determined to be significant and severe, it is still unknown which of those discrepancies are due to a specific learning disability, low motivation, cultural differences, or educational disadvantage. Consequently, additional criteria are needed to differentitate the learning disabled population from other low-achieving groups. Second, most psychometric instruments used in discrepancy models are inappropriate with populations of Limited English Speaking (LES) or Non-English Speaking (NES) individuals because these tests lack non-English versions and norms. Thus, the validity of the discrepancy criterion must be evaluated for these subgroups. Third, other components of the existing LD definition have not been operationalized. For example, although the "exclusion clause" of the LD definition emphasizes that the academic discrepancy may not be due primarily to such causes as severe emotional disturbance, environmental or economic disadvantage, cultural/language differences, etc., no empirically based examplars are available to indicate how these criteria should be operationalized. In summary, it is evident that a major need exists to empirically validate a set of factors to be used in conjunction with the discrepancy factors in making LD identification decisions.

5. Eligibility scores for student admission to LD programs must be established. Prior to PL 94-142 and the institutionalization of specific learning disabilities as a handicapping condition, the LD population was composed of a small, clinically defined group. However, as programs for the learning disabled have grown in popularity in recent years, the number of students being referred for LD services has grown even more rapidly. Because of limited resources decisions are now being made to establish cutoff levels for determining who should be eligible for such services and who should not. In the past, clinical judgment constituted a major identification strategy for LD populations. However, many of these clinical decision-making practices failed to systematically consider the values of key stakeholder groups (e.g., college administrators, LD specialists, Department of Finance personnel, etc.) in deciding the minimum eligibility criteria. The issue of setting eligibility cutoff scores is as central to the entire identification issue as the specification of which test instruments to use, for example. Appropriate cutoff levels can be established by using decision-making methodologies that explicitly consider the values, concerns, motives, and interests of the key stakeholders involved. In addition, as these procedures are based on both clinical and statistical criteria they can take into consideration the rates of correct and incorrect classifications, and assess the positive and negative consequences of any decision regarding cutoff levels.

6. Empirically validated data are lacking on existing identification procedures. To date, no comprehensive studies have been conducted within the learning disability field to validate current identification practices. Until such practices are submitted to a rigorous empirical validation study, the field will be unable to justify decision-making policy or intervention

practices. It is imperative that any proposed eligibility criteria be thoroughly tested to determine their efficacy for field application. It should be noted that numerous studies have essentially provided an "armchair" analysis of one or two components of the eligibility criteria and procedures. Just as often, what was offered as promising, innovative identification procedures turned out to be far less, and actually resulted in more confusion since they promised much but produced little of significant utility. Until the problems previously mentioned have been collectively resolved and validated in an empirical study, the proposed identification process and procedures are "strawmen"; they represent best guesses and the practical consequences of a given model remain unknown. The California Community College System is in a unique position of being able to systematically gather the data necessary for validating its selection procedures and empirically determining the requirements and consequences of their implementation. If this opportunity is ignored, the related problems will only increase in number, range, and severity.

Management Plan

The challenge of developing minimum eligibility criteria for identifying LD adults in community college programs is great, inasmuch as these criteria must be not only conceptually sound but also empirically based. Given the complex nature of this undertaking, the Steering Committee took considerable care in organizing the personnel and management aspects of the project. The following two criteria were considered critical in meeting the Consortium's goals.

First, given the array of issues and decision points associated with the development of an LD identification procedure, it was necessary to systematically solicit a broad base of input from key individuals, professional organizations, and state agencies. The Management Chart in Figure __ depicts how each project goal was supported by a Task Force(s) to facilitate the research and decision-making process. In addition, a State Advisory Committee was established to provide input to the Steering Committee about the overall direction and activities of the project. Individuals participating on the task forces and the advisory committee were selected to provide a broad base of representativeness and expertise to the project.

Second, numerous issues requiring technical expertise were encountered throughout the project. From the beginning, the Steering Committee has been committed to accessing experts with national/international reputations to consultants who provided such expertise is provided in the Management Chart (see Figure A). No activity conducted in the learning disability field to date has utilized expertise equal to that tapped for the present project. In addition to providing input that was on the "cutting edge," these experts brought another very critical dimension to the efforts of the project - objectivity. That is, none of the experts had a vested interest in the direction or outcome of the project. This was felt to be critical since the importance of adhering to a rigorous set of scientifically defensible research procedures was deemed so great.

In summary, the Steering Committee views as one of the greatest strenghts of the project the broad base of input received from key experts and practitioners throughout the course of the research. In addition to an effective, and very necessary system of checks and balances, a process has been developed whereby some of the best thinking in learning disabilities and related fields was tapped to insure successful goal attainment.

Project Goals and Outcomes

Given the problems outlined earlier, four goals were specified as being necessary for the development of minimum eligibility criteria.

Although accomplishing these goals might well require a project spanning several years, time and budget constraints dictated that many of the related activities be completed within one year. The following information is presented to provide a brief description of project activities toward meeting the specified goals. This description includes a statement of the four goals, and the intended outcome, procedures, major findings, and proposed actions of each goal.

GOAL 1 To specify educationally relevant attributes and technically adequate instruments that differentiate learning disabled adults from other low achieving adults.

Intended Outcome: To specify a list of attributes which best differentiate LD from low-achieving students in community colleges.

Procedures: Three activities were completed to specify the educationally relevant attributes that differentiate the LD from other low achievers. First, a thorough search of the research literature was conducted to identify potential differentiating attributes. In addition, LD specialists on Task Force I were asked to specify differentiating attributes based on their experience with LD adults. Second, the attributes identified in the first step were honed to 116 and organized into a survey. Third, the survey of attributes was completed by 281 professionals who had teaching or research experience with learning disabled and other low achieving adults.

Major findings:

1. 116 specific attributes were identified through a search of the research literature and surveying LD practitioners.

2. Based on the judgments of 281 practitioners, researchers and university/college faculty, 55 items were statistically identified as distinguishing the LD from other low-achievers.

3. Ten of the 55 items focused on reading related skills e.g., omits words while reading, loses place while reading, has a discrepancy between word attack and passage comprehension, etc.

4. Eight of the 55 items were concerned with written communication skills (e.g., difficulty in detecting errors in their writing, difficulty in clearly expressing ideas in writing, has poor spelling ability).

5. The item content of eight of the 55 items focused on the cause of the poor school performance e.g., genetic predisposition, head trauma, memory problem, etc.

6. On seven of the 55 items statistically significant differences were determined among the groups providing the ratings (e.g., LD practitioner, resource teacher, adminstrator, psychologist, etc.)

Proposed action:

1. Review the 55 items for representativeness and for clarity of content.
2. Determine how these items might be best applied in differentiating the LD from other low-achivers.
3. Develop normative data with community college LD and non-LD populations for the differentiating attributes.

Intended outcome: Specify a set of reliable and valid instruments that can be used to operationally measure each area of learning disability.

Procedures. To specify the technical adequacy (i.e., appropriate normative data, reliability coefficients, derived scores, and validity indices) of assessment instruments used with LD adults, three experts in the area of assessment (See Management Chart, Appendix B) were asked to evaluate the test instruments most frequently used or recommended for this population (Ostertag and Baker, 1982). This group of tests included the following: the Key Math Diagnostic Arithmetic Test, the Peabody Individual Achievement Test the revised edition of the Peabody Picture Vocabulary Test, Raven's Progressive Matrices, the Wechsler Adult Intelligence Scale, the revised edition (1981) of the Wechsler Adult Intelligence Scale, the Woodcock-Johnson Psycho-Educational Battery, Parts I and II, the Woodcock Reading Mastery Test, and the Wide Range Achievement Test. A standard format was provided for these evaluations.

As one might surmise, the initial listing of nine instruments is not extensive neither in terms of the number of tests for any one area (e.g., written expression) nor of evaluating all seven deficit areas associated with a learning disability. However, our search of available assessment instruments revealed a lack of adequate measures of such skill areas as oral expression or listening comprehension.

Major findings:

1. Three measures of academic potential were judged to be technically adequate: Wechsler Adult Intelligence Scale, the Wechsler Adult Intelligence Scale - Revised, and the Woodcock-Johnson Psycho-educational Battery, Part I. These measures could be applied in the discrepancy model.
2. Two measures of academic achievement were judged technically adequate for the discrepancy method: the Woodcock-Johnson Psycho-educational Battery, Part II and the Wide Range Achievement Test.

Proposed action:

1. Review these tests with practitioners for comments.
2. Elicit names of tests for which appropriate technical data is lacking e.g., normative data, reliability coefficients, but for which there is support for obtaining this information
3. Determine plan for obtaining needed technical data relevant to the assessment instruments and their use in the discrepancy method.

GOAL 2 To develop appropriate discrepancy/selection models to verify the presence of learning disabilities in community college students.

Intended outcome: To specify a discrepancy/selection model which is appropriate for community college settings.

Procedures. To successfully address this goal it was necessary to develop a plan of action that permitted a thorough analysis and subsequent choice of discrepancy models most appropriate for application in community college settings. First, a comprehensive review and analysis of existing statistical discrepancy models was completed through search of the published literature and querying nationally recognized experts from the areas of educational and psychological measurement for their analysis of different discrepancy models. Second, the University of Kansas Research Institute staff prepared a lengthy paper which evaluated various discrepancy formulas in terms of their feasibility and technical characteristics. This paper also included the measurement experts' recommendations. Third, the paper was submitted to Task Force II for review.

Major findings:

As a result of Task Force II, the following statements were endorsed by Task Force II:

1. A discrepancy model was one appropriate eligibility criterion for LD programs.

2. An Estimated Regressed True Score Method for calculating a discrepancy score was endorsed.

3. Alternative discrepancy methods need to be evaluated along with the preferred option.

4. The need for a selection model and criteria for those students for whom it is invalid to apply a discrepancy model e.g., ESL, LES-NES, was confirmed.

Proposed action:

1. To obtain the necessary technical data for those aptitude-achievement test combinations to be used in the discrepancy model e.g., correlation coefficients.

2. To obtain the necessary normative data on those factors in a selection model so that the learning disabled and non-learning disabled populations demonstrate observable, distinct, and reliable performance differences.

GOAL 3 To establish the magnitude of discrepancy necessary to verify a learning disability among community College students.

Intended outcome: To provide a level of student performance from which the verification of LD can be made on a consistent and agreed upon basis.

Procedures. The following activities were undertaken to address this goal. First, the Steering Committee identified the state agencies, organizations, and groups who might possess significant values about eligibility standards, i.e., a cutoff score for the community college LD program. Organizations expressing similar concerns were combined into one stakeholder group. A total of seven unique groups were identified: (a) Underachieving students; (b) LD specialists; (c) Legislative group; (d) College administrators; (e) State agencies and administration; (f) Non-LD college instructors; and (g)

Ancillary services group. A representative of each stakeholder group was asked to participate in an elicitation session in March under the direction of Dr. Detlof von Winterfeldt. These participants composed Task Force III.

Second, prior to the March elicitation session, Dr. von Winderfeldt prepared a set of Multiattribute Utility Questionnaires for use by Task Force III. The questionnaires, which were constructed from data provided by professionals in the field throughout the State of California, had both an instructional function concerning classification decisions and a data-gathering purpose in determining an eligibility criterion.

Third, at the March elicitation activity, Task Force III members (a) explicitly stated their values concerning the classification problem; (b) evaluated and ranked the desirability and undesirability of classification decisions for each value; and (c) weighted the relative importance of their values. Central to these classification activities is the existence of both correct decisions (e.g., classifying a "true" LD student as LD) and incorrect decisions (e.g., classifying a "true" non-LD student as LD).

Representatives (a total of 17) for each of the seven stakeholder groups were in attendance at the Task Force meeting on March 24.

The computations from the Task Force activities provide the State of California with a criterion likelihood ratio that must be met in order for a particular discrepancy score to indicate a possible specific learning disability. This criterion was developed by formally considering the values thought to be most important by representatives of seven stakeholder groups and the positive and negative consequences of misclassifying students as either LD or non-LD.

It should be noted that the present project represents the first reported instance in which a formalized decision-theoretic approach was followed to establish a criterion score for eligibility in LD programs. In short, this procedure quantified the potentially competing values of various stakeholder groups and the utility of decisions based on those values. Such an approach was thought necessary for numerous reasons including the need to provide consistent, clearly defined criteria for the many community college settings in California.

Major findings:

1. Based on the expressed values of stakeholder groups, a criterion odds ratio of 13.64 to 1 should be met to identify a learning disabled student. Or in other words, the cutoff score should be such that 93% of the time the student is correctly identified as learning disabled.
2. The non-LD instructors desired the most lenient cutoff score; the Department of Finance desire the strictest cutoff score.
3. This odds ratio can be applied to any criteria on which eligibility is determined and therefore is not limited to just an aptitude-achievement discrepancy model.

Proposed action:

1. For each of aptitude-achievement test pairings in the discrepancy model, obtain a distribution of discrepancy scores from which a cutoff level can be determined.
2. Evaluate this cutoff level for evidence of population bias e.g., differences in culture, sex, age, and race.
3. To agree upon an incidence level for LD in the population which can then be used in determining the cutoff score in the eligibility criteria. (This is sometimes described as a prior probability or marginal probability.)

GOAL 4 To complete a pilot study of the proposed discrepancy and selection models.

Intended outcome: To provide descriptive information on the results of applying the discrepancy and selection model.

Procedures. The original plan suggested that a validation study be completed during the current academic year. However, as the complexities of the issues to be dealt with became evident, the Steering Committee recommended that further study be conducted rather than a full-scale validation effort. This recommendation was made to insure that efforts to accomplish goals 1, 2, and 3 would be sufficiently comprehensive and rigorous.

This study began in mid-February with an April 15th deadline for all data collection. Twenty-one community college campuses throughout the State of California participated.

The study of the discrepancy model assessed the impact of applying a discrepancy model and involved the assessment of both LD and other randomly selected students. These subjects were evaluated on measures of academic achievement as well as intellectual ability. The subsequent data analysis considered not only the specific scores obtained by the different samples, it also evaluated the process and procedures used to obtain the data (e.g., Did "random" selection prove to be a difficult procedure? Were nonhandicapped students willing to participate as volunteers? How much time was required to complete the procedures?). Evaluation of specific discrepancy scores required calculating discrepancy scores using the True-Score Regression formula (recommended by Task Force II) and other discrepancy methods and subsequently comparing the resulting scores.

The selection model study also entailed evaluating the responses of specific criterion groups of learning disabled and other low achieving students to survey instruments and evaluating the process and procedures used to obtain the data. The results from these surveys plus the information obtained in Goal 1 concerning LD attributes provide a data base for operationalizing other criteria which must be considered in verifying a learning disability.

Major findings of discrepancy model:

1. Using a 10% incidence level which is a very lenient criterion and any of four discrepancy methods, the percent of currently identified LDs who would qualify in reading is 35% (17) in written expression is 46% (23) and in arithmetic is 26% (13). The numbers in parenthesis correspond to the actual number of students in the study who met the criterion.
2. The Regressed Score method (Cone and Wilson, 1981) identified the most students in any area.
3. For comparison, using the same formula, incidence level, and achievement areas, the following percent of randomly selected students met the criteria: 14%, 17% and 23% respectively.
4. The mean IQ of the randomly selected students was 108 (SD = 14.4). On the three areas of achievement, the scores were up to four points above the population mean.
5. The mean IQ of the LD students was 83 (SD = 9.3). The mean scores on the three achievement areas were approximately 81 in reading, 80 in written expression, and 87 in math.
6. The profile of the LD population appears to be that of the low achieving student who is achieving at his/her expected achievement level.

Proposed action:

1. Determine additional tests on which normative data is requested.
2. Obtain the data necessary for calculating required correlation coefficients based on representative samples.

Major findings of selection model:

1. For the populations studied, the LD and low-achieving (LA) students are very alike.
2. The differences between the LD and LA populations centered on 12 variables which clustered into the following five areas:
    a. Academic performance: LAs were comparatively better than LD on the following: Woodcock-Johnson Reading and Written Expression cluster scores, ability to acquire test taking skills, ability to work independently, ability to participate in discussions, ability to learn from printed material, ability to learn reading and spelling, and reading recognition skills.
    b. Self-help: LD seek tutorial assistance more often than do LA. Eligibility in an LD program may be dependent on how this action is perceived by the LD specialist.
    c. Sex: Fewer females were classified as LD than males.
    d. Exclusionary factor: LAs change junior high and high schools more frequently than do LD. This may reduce the possibility of an LA being considered LD.
    e. Social activity: LAs are more active in social organizations than are the LD students.
3. Total sample size was 49 of which 15 were LA and 34 were identified as LD.

4. Only 6 students indicated ESL.
5. The average age of all students was 27 years. The modal age was 20 (8 students), and the median age was 23.5.

Comparison of Successful and Unsuccessful LD students:

1. Of the 34 LD students in the study, 18 were coded as successful, and 16 were coded unsuccessful.
2. Group differences were of degree and not qualitative on the 16 significant variables.
3. Unsuccessful LD students uniformly showed greater decrements of performance on such items as ability to stay on task, taking notes in class, missing classes, completing homework assignments, spelling, identifying realistic career goals, and paying attention in class.

Proposed action:

1. Determine the utility of the successful and unsuccessful LD student dichotomy for normative and validation study.
2. Determine the representativeness of the significant survey items as they might be used in eligibility criteria (e.g., How does this information help with the ESL-LD student?)
3. Determine plan for developing assessment instrument incorporating knowledge obtained from pilot study data.
4. Determine the utility of continuing to research subpopulation bias in selection procedures.
5. Proceed with the belief that an age related variable will be needed for norming instruments and calculating discrepancy score distributions.

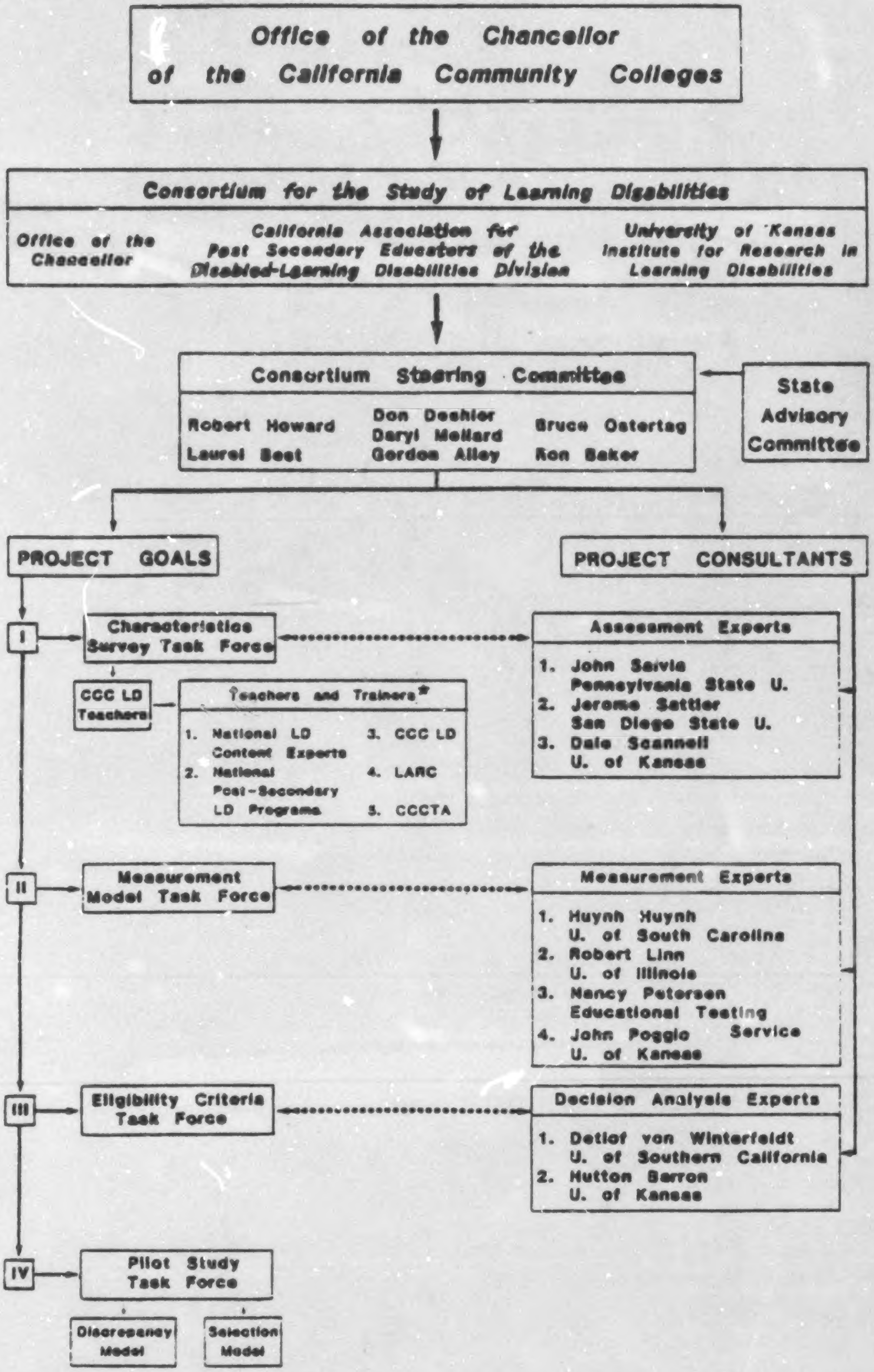
Conclusion

As a result of the Consortium's activities during the past year, a solid foundation has been laid in the State of California for gaining resolution of a major educational problem, that is, the development of minimum eligibility criteria and procedures for identifying learning disabled adults in community college programs. The necessary steps have been taken from both theoretical and empirical standpoints, to systematically design such standards. To validate the procedures developed thus far and to gather the information necessary for implementing the procedures, it is recommended that the California community college system complete a norming and validation study. Such a study of both the discrepancy and selection models will provide the state with the necessary data for implementing eligibility criteria and procedures. Hence the standardization of eligibility criteria throughout the state is ensured.

While several states and school districts throughout the country profess to having "standardized" procedures for LD identification, none has validated their procedures empirically. Consequently, problems of inappropriate placements, inflated enrollments, and discrimination abound. In contrast, the data base provided through a norming and validation study will allow the State of California not only to be responsive to the needs of learning disabled adults but also to become more responsible in the discharge of its overall educational mission to all community college students.

BEST COPY AVAILABLE

- H


Office of the Chancellor of the California Community Colleges
Consortium for the Study of Learning Disabilities
Office of the Chancellor
California Association for Post Secondary Educators of the Disabled-Learning Disabilities Division
University of Kansas Institute for Research in Learning Disabilities
Consortium Steering Committee
Robert Howard
Laurel Best
Don Deshler
Daryl Mellard
Gordon Alley
Bruce Ostertag
Ron Baker
State Advisory Committee
PROJECT GOALS
PROJECT CONSULTANTS
I
Characteristics Survey Task Force
CCC LD Teachers
Teachers and Trainers*
1. National LD Content Experts
2. National Post-Secondary LD Programs
3. CCC LD
4. LARC
5. CCCTA
Assessment Experts
1. John Salvia Pennsylvania State U.
2. Jerome Sattler San Diego State U.
3. Dale Scannell U. of Kansas
II
Measurement Model Task Force
Measurement Experts
1. Huynh Huynh U. of South Carolina
2. Robert Linn U. of Illinois
3. Nancy Petersen Educational Testing Service
4. John Poggio U. of Kansas
III
Eligibility Criteria Task Force
Decision Analysis Experts
1. Detlof von Winterfeldt U. of Southern California
2. Hutton Barron U. of Kansas
IV
Pilot Study Task Force
Discrepancy Model
Selection Model

* CCC LD: California Community College Learning Disabilities Specialists
LARC: Learning Resources Association of the California Community Colleges
CCCTA: California Community College Tutorial Association

November 19, 1982

RECOMMENDATIONS

As a result of the Consortium's activities during the past 10 months, a solid foundation has been laid in the state of California for resolving a major educational problem, that is, the development of minimum eligibility criteria and procedures for identifying learning disabled adults in community college programs. From both theoretical and empirical standpoints the necessary steps have been taken to systematically design such standards. To validate the procedures developed to date and to gather the information necessary for implementing the procedures, it is recommended that the California community college system complete a comprehensive follow-up study.

Such a follow-up study of both the discrepancy and selection models will provide the state with the necessary data for empirically validating the minimum eligibility criteria and procedures. Hence the standardization of eligibility criteria throughout the state is insured.

While several states and school districts profess to having "standardized" procedures for LD identification, none has validated their procedures empirically. Consequently, problems related to inappropriate placements, inflated enrollments, and discrimination abound. In contrast, the data base provided through a follow-up study will allow the state of California not only to be responsive to the needs of learning disabled adults but also to become more responsive in the discharge of its overall educational mission to all community college students. The following section provides a plan of action for accomplishing the validation phase of this effort. It is estimated that approximately 24 months would be required to complete the workscope.

Goal 1

To collect the technical data required for implementing the statistical discrepancy model in California Community Colleges.

Objective 1 To confirm the assessment instruments to be used to operationalize the discrepancy model.

1.1 To solicit judgments from practitioners about the appropriateness of the assessment instruments judged to be technically adequate by assessment experts.

Objective 2 To obtain normative data on technically adequate instruments that do not have appropriate norms for CCC.

Objective 3 To determine the correlations between all aptitude and achievement measures specified in Objective 1.0 and 2.0.

Objective 4 To determine discrepancy scores and academic aptitudes of LD students.

Goal 2

To collect normative and validation data for the Community College LD Attribute Survey and other criteria (e.g., "Exclusion clause")

Objective 1 Based on results of Survey 2 and the Pilot Study, confirm the behavioral domains for validity in identifying the LD in CC.

Objective 2 To clarify the exclusion clause for implementation as a part of the eligibility criteria.

NOTE: Those practitioners completing Objective 1 could also complete this objective.

Objective 3 To determine the most appropriate scaling procedures (paired-comparison, Likert-type, Thurstone equal-appearing interval scale, rating scale, etc.) and format (self-report, interview, etc.) for obtaining information.

Objecitve 4 To develop content of LD Attribute Survey and exclusion clause criteria.

Objective 5 To gather psychometric data on the LD Attribute Survey and other selection criteria on a nonhandicapped and LD sample.

Goal 3

To field test the identification procedures.

Objective 1.0 To define a flow chart (including a narrative description of the basic procedures) required for satisfying minimum eligibility criteria for depicting events and identifying LD community college students.

1.1 Present an elementary draft of the flow-chart to fifty practitioners (including LD specialists, psychologists, Deans of Students Services, State Advisory Committee, etc.) for review and comment.

1.1.1 Analyze comments and modify flow chart procedures.

Objective 2.0 To implement the identification model (flow chart) using different decision models (e.g., regression equation, multiple cut-off, decision theoretic, Bayesian).

2.1 To evaluate alternative selection models specified in the literature.

2.1.1 Have measurement experts review, evaluate, and recommend the most appropriate selection models for application with a community college population.

Objective 3.0 To specify the parameters for conducting the field test.

3.1 Identify 30 community college sites (or sufficient sites to make data available on 250 students) to participate in the field test.

3.2 Prepare materials for the field test.

3.3 Train site coordinators in application of the field test procedures.

3.4 Develop evaluation procedures for monitoring the implementation of the identification model.

3.5 Collect and analyze data.

Objective 4.0 To make revisions as specified by field-test participants.

4.1 Convene a Task Force of Practitioners to review the field test data to make decisions for the final identification model.

MASTER PLAN FOR DATA COLLECTIONS

A significant amount of data will be required to successfully complete the proposed follow-up study. Given the budget constraints and the heavy loads carried by community college professionals associated with learning disabled students, it is imperative to design an effective set of procedures/policies to insure maximum data return for minimum cost and intrusion to the system. To have sufficient power and validity all data collected must come from representative samples.

Thus, the success of this phase of the project is contingent upon a strong endorsement of the goals and objectives of the project from the Chancellor's Office. Such endorsement would include: (a) a letter from the Chancellor to community college presidents in participating sites requiring their participation and full cooperation; (b) the Chancellor requiring participating schools to designate someone to assume the responsibility for coordinating and insuring completion of all research activities on campus; and (c) bringing the "powers" of the Chancellor's Office to bear on schools failing to comply with the data collections needs of the study. Without this support to insure collection of necessary data, the goals of the validation study cannot be reached.

ERIC Clearinghouse for Junior Colleges
AUG 16 1985